

LARGE DEVIATION TYPE ESTIMATES FOR RANDOM COCYCLES

PEDRO DUARTE AND SILVIUS KLEIN

ABSTRACT. In this paper we prove the continuity of all Lyapunov exponents, as well as the continuity of the Oseledets decomposition, for a class of irreducible cocycles over strongly mixing Markov shifts. Moreover, gaps in the Lyapunov spectrum lead to a Hölder modulus of continuity for these quantities. This result is an application of the abstract continuity theorems obtained in [7], and generalizes a theorem of E. Le Page on the Hölder continuity of the maximal LE for one-parameter families of strongly irreducible and contracting cocycles over a Bernoulli shift.

This is a draft of a chapter in our forthcoming research monograph [7].

1. INTRODUCTION AND STATEMENTS

We define the class of random cocycles over Markov shifts and describe our assumptions on them. We then formulate the main statements, and sketch the argument for proving large deviation type estimates. Finally we relate our findings to other results for similar models.

1.1. Description of the model. Let Σ be a compact metric space and \mathcal{F} its Borel σ -field.

Definition 1.1. A Markov kernel is a function $K : \Sigma \times \mathcal{F} \rightarrow [0, 1]$ such that

- (1) for every $x \in \Sigma$, $A \mapsto K(x, A)$ is a probability measure in Σ , also denoted by K_x ,
- (2) for every $A \in \mathcal{F}$, the function $x \mapsto K(x, A)$ is \mathcal{F} -measurable.

The iterated Markov kernels are defined recursively, setting

- (a) $K^1 = K$,
- (b) $K^{n+1}(x, A) = \int_{\Sigma} K^n(y, A) K(x, dy)$, for all $n \geq 1$.

Each power K^n is itself a Markov kernel on (Σ, \mathcal{F}) .

A probability measure μ on (Σ, \mathcal{F}) is called *K-stationary* if for all $A \in \mathcal{F}$,

$$\mu(A) = \int K(x, A) \mu(dx) .$$

A set $A \in \mathcal{F}$ is said to be *K-invariant* when $K(x, A) = 1$ for all $x \in A$ and $K(x, A) = 0$ for all $x \in X \setminus A$. A K -stationary measure μ is called *ergodic* when there is no K -invariant set $A \in \mathcal{F}$ such that $0 < \mu(A) < 1$. As usual, ergodic measures are the extremal points in the convex set of K -stationary measures.

Definition 1.2. A Markov system is a pair (K, μ) , where K is a Markov kernel on (Σ, \mathcal{F}) and μ is a K -stationary probability measure.

Let (K, μ) be a Markov system. There is a canonical construction, due to Kolmogorov, of a probability space $(X, \mathcal{F}, \mathbb{P}_\mu)$ and a Markov stochastic process $\{e_n : X \rightarrow \Sigma\}_{n \geq 0}$ with initial distribution μ and transition kernel K , i.e., for all $x \in \Sigma$ and $A \in \mathcal{F}$,

- (1) $\mathbb{P}_\mu[e_0 \in A] = \mu(A)$,
- (2) $\mathbb{P}_\mu[e_n \in A | e_{n-1} = x] = K(x, A)$.

We briefly outline this construction. Elements in Σ are called *states*. Consider the space $X^+ = \Sigma^\mathbb{N}$ of *state sequences* $x = (x_n)_{n \in \mathbb{N}}$, with $x_n \in \Sigma$ for all $n \in \mathbb{N}$, and let \mathcal{F}^+ be the product σ -field $\mathcal{F}^+ = \mathcal{F}^\mathbb{N}$ generated by the \mathcal{F} -cylinders, i.e., generated by sets of the form

$$C(A_0, \dots, A_m) := \{x \in X^+ : x_j \in A_j, \text{ for } 0 \leq j \leq m\},$$

where $A_0, \dots, A_m \in \mathcal{F}$ are measurable sets. The (topological) product space X^+ is compact and metrizable. The σ -field \mathcal{F}^+ coincides with the Borel σ -field of the compact space X^+ .

Definition 1.3. Given any probability measure θ on (Σ, \mathcal{F}) , the following expression determines a pre-measure

$$\mathbb{P}_\theta^+[C(A_0, \dots, A_m)] := \int_{A_m} \cdots \int_{A_0} \theta(dx_0) \prod_{j=1}^m K(x_{j-1}, dx_j)$$

on the semi-algebra of \mathcal{F} -cylinders. By Carathéodory's extension theorem this pre-measure extends to a unique probability measure \mathbb{P}_θ^+ on (X^+, \mathcal{F}^+) .

It follows from this definition that the sequence of random variables $e_n : X^+ \rightarrow \Sigma$, defined by $e_n(x) := x_n$ for $x = (x_n)_{n \in \mathbb{N}}$, is a Markov chain with initial distribution θ and transition kernel K w.r.t. the probability space $(X^+, \mathcal{F}^+, \mathbb{P}_\theta^+)$. It also follows that the process $\{e_n\}_{n \geq 0}$ is stationary w.r.t. $(X, \mathcal{F}^+, \mathbb{P}_\theta^+)$ if and only if θ is a K -stationary measure.

Consider now the space $X = \Sigma^\mathbb{Z}$ of bi-infinite *state sequences* $x = (x_n)_{n \in \mathbb{Z}}$, with $x_n \in \Sigma$ for all $n \in \mathbb{Z}$, and let \mathcal{F} be the product σ -field $\mathcal{F} = \mathcal{F}^\mathbb{Z}$ generated by the \mathcal{F} -cylinders in X . Again the topological product space X is both metrizable and compact, and the σ -field \mathcal{F} is

the Borel σ -field on the compact metric space X . There is a canonical projection $\pi : X \rightarrow X^+$, defined by $\pi(x_n)_{n \in \mathbb{Z}} = (x_n)_{n \in \mathbb{N}}$, relating these two spaces.

Markov systems are probabilistic evolutionary models, which can also be studied in dynamical terms. For that we introduce the *shift mappings*.

Definition 1.4. *The one-sided shift is the map $T : X^+ \rightarrow X^+$, $T(x_n)_{n \geq 0} = (x_{n+1})_{n \geq 0}$, while the two-sided shift is the map $T : X \rightarrow X$, $T(x_n)_{n \in \mathbb{Z}} = (x_{n+1})_{n \in \mathbb{Z}}$.*

The map $T : X^+ \rightarrow X^+$ is continuous, and hence \mathcal{F}^+ -measurable. It also preserves the measure \mathbb{P}_μ^+ , i.e., $T_*\mathbb{P}_\mu^+ = \mathbb{P}_\mu^+$. Moreover, the Markov process $\{e_n\}_{n \geq 0}$ on $(X^+, \mathcal{F}^+, \mathbb{P}_\mu^+)$ is dynamically generated by the observable e_0 in the sense that $e_n = e_0 \circ T^n$, for all $n \geq 0$.

The two-sided-shift $T : X \rightarrow X$ is a homeomorphism, and hence \mathcal{F} -bimeasurable. The projection $\pi : X \rightarrow X^+$ semi-conjugates the two shifts. The two-sided-shift is the *natural extension* of the one-sided-shift. According to this construction (see [17]), there is a unique probability measure \mathbb{P}_μ on (X, \mathcal{F}) such that $T_*\mathbb{P}_\mu = \mathbb{P}_\mu$ and $\pi_*\mathbb{P}_\mu = \mathbb{P}_\mu^+$. We will refer to the measures \mathbb{P}_μ^+ and \mathbb{P}_μ as the *Kolmogorov extensions* of the Markov system (K, μ) .

Definition 1.5. *Given a Markov system (K, μ) let \mathbb{P}_μ be the Kolmogorov extension of (K, μ) on $X = \Sigma^\mathbb{Z}$. The dynamical system (X, \mathbb{P}_μ, T) is called a Markov shift.*

Let $(L^\infty(\Sigma), \|\cdot\|_\infty)$ denote the Banach algebra of complex bounded \mathcal{F} -measurable functions with the sup norm $\|f\|_\infty = \sup_{x \in \Sigma} |f(x)|$. The following concept corresponds to condition (A1) in [1].

Definition 1.6. *We say that a Markov system (K, μ) is strongly mixing if there are constants $C > 0$ and $0 < \rho < 1$ such that for every $f \in L^\infty(\Sigma)$, all $x \in \Sigma$ and $n \in \mathbb{N}$,*

$$\left| \int_\Sigma f(y) K^n(x, dy) - \int_\Sigma f(y) \mu(dy) \right| \leq C \rho^n \|f\|_\infty.$$

It follows from this definition that,

Proposition 1.1. *If the Markov system (K, μ) is strongly mixing then the Markov shift (X, \mathbb{P}_μ, T) is a mixing dynamical system.*

Proof. Consider a bounded measurable observable $f : X \rightarrow \mathbb{R}$ depending only on the coordinates x_0, \dots, x_p , and write $f(x) = f(x_0, \dots, x_p)$. Let $g(x) = g(x_{-q}, \dots, x_{-1})$ be another bounded measurable observable depending only on the coordinates x_{-q}, \dots, x_{-1} with $q \in \mathbb{N}$. Denote

by $\{e_n\}_{n \in \mathbb{Z}}$ the Markov process on (X, \mathbb{P}_μ) with common distribution μ and transition kernel K . By the strong mixing property

$$\mathbb{E}_{x_0}[f(e_n, \dots, e_{n+p})] = \int_{\Sigma} \cdots \int_{\Sigma} f(x_n, \dots, x_{n+p}) K^n(x_0, dx_n) \prod_{j=n}^{n+p-1} K(x_j, dx_{j+1})$$

converges uniformly (in x_0) to

$$\int_{\Sigma} \cdots \int_{\Sigma} f(x_n, \dots, x_{n+p}) \mu(dx_n) \prod_{j=n}^{n+p-1} K(x_j, dx_{j+1}) = \mathbb{E}_\mu(f) .$$

Hence

$$\begin{aligned} \mathbb{E}_\mu[(f \circ T^n) g] &= \mathbb{E}_\mu[g(e_{-q}, \dots, e_{-1}) f(e_n, \dots, e_{n+p})] \\ &= \int_{\Sigma} \cdots \int_{\Sigma} g(x_{-q}, \dots, x_{-1}) \mathbb{E}_{x_0}[f(e_n, \dots, e_{n+p})] \mu(dx_{-q}) \prod_{j=-q}^{-1} K(x_j, dx_{j+1}) \end{aligned}$$

converges to

$$\mathbb{E}_\mu(f) \int_{\Sigma} \cdots \int_{\Sigma} g(x_{-q}, \dots, x_{-1}) \mu(dx_{-q}) \prod_{j=-q}^{-1} K(x_j, dx_{j+1}) = \mathbb{E}_\mu(f) \mathbb{E}_\mu(g) .$$

The mixing property of the shift (X, \mathbb{P}_μ, T) follows applying the previous argument to the indicator functions of any two cylinders, and noting that the σ -algebra of cylinders generates the Borel σ -field of X . \square

Examples of strongly mixing Markov systems arise naturally from Markov kernels satisfying the *Doeblin condition* (see [3]). We say that K satisfies the Doeblin condition if there is a positive finite measure ρ on (Σ, \mathcal{F}) and some $\varepsilon > 0$ such that for all $x \in \Sigma$ and $A \in \mathcal{F}$,

$$K(x, A) \geq 1 - \varepsilon \quad \Rightarrow \quad \rho(A) \geq \varepsilon .$$

Given $A \in \mathcal{F}$, define

$$L^\infty(A) := \{ f \in L^\infty(\Sigma) : f|_{\Sigma \setminus A} \equiv 0 \} ,$$

which is a closed Banach sub-algebra of $(L^\infty(\Sigma), \|\cdot\|_\infty)$.

Proposition 1.2. *Let (Σ, K) be a Markov system. If K satisfies the Doeblin condition then there are sets $\Sigma_1, \dots, \Sigma_m$ in \mathcal{F} and probability measures ν_1, \dots, ν_m on Σ such that for all $i, j = 1, \dots, m$,*

- (1) $\Sigma_i \cap \Sigma_j = \emptyset$ when $i \neq j$,
- (2) Σ_i is K -forward invariant, i.e., $K(x, \Sigma_i) = 1$ for $x \in \Sigma_i$,
- (3) ν_i is K -stationary and ergodic with $\nu_i(\Sigma_j) = \delta_{ij}$,
- (4) $\lim_{n \rightarrow +\infty} K^n(x, \Sigma_1 \cup \dots \cup \Sigma_m) = 1$, with geometric uniform speed of convergence, for all $x \in \Sigma$,

(5) $\nu(\Sigma_1 \cup \dots \cup \Sigma_m) = 1$, for every K -stationary probability ν .

Moreover, for every $1 \leq i \leq m$ there is an integer $p_i \in \mathbb{N}$ and measurable sets $\Sigma_{i,1}, \dots, \Sigma_{i,p_i} \in \mathcal{F}$ such that

- (1) $\{\Sigma_{i,1}, \dots, \Sigma_{i,p_i}\}$ is a partition of Σ_i ,
- (2) $K(x, \Sigma_{i,j+1}) = 1$ for $x \in \Sigma_{i,j}$ and $1 \leq j \leq p_i$, with $\Sigma_{i,p_i+1} = \Sigma_{i,1}$,
- (3) $(\Sigma_{i,j}, K^{p_i})$ is strongly mixing for all $1 \leq j \leq p_i$.

Proof. See [3, section V-5]. \square

Let (K, μ) be a Markov system. We introduce a space of measurable functions $A : \Sigma \times \Sigma \rightarrow \text{GL}(m, \mathbb{R})$.

Definition 1.7. The space $\mathcal{B}_m^\infty(K)$ consists of all functions $A : \Sigma \times \Sigma \rightarrow \text{GL}(m, \mathbb{R})$ such that A and A^{-1} are both measurable and uniformly bounded. On this space we consider the metric $d_\infty(A, B) = \|A - B\|_\infty$.

Definition 1.8. The function $A \in \mathcal{B}_m^\infty(K)$ determines a linear cocycle $F_A : X \times \mathbb{R}^m \rightarrow X \times \mathbb{R}^m$ over the Markov shift (X, \mathbb{P}_μ, T) , defined by

$$F_A(x, v) := (Tx, A(x)v),$$

where we identify A with the function $A : X \rightarrow \text{GL}(m, \mathbb{R})$, $A(x) := A(x_0, x_1)$, for $x = (x_n)_{n \in \mathbb{Z}} \in X$.

The iterates of F_A are the maps $F_A^n : X \times \mathbb{R}^m \rightarrow X \times \mathbb{R}^m$,

$$F_A^n(x, v) = (T^n x, A^{(n)}(x)v),$$

with $A^{(n)} : X \rightarrow \text{GL}(m, \mathbb{R})$ defined for all $x = (x_n)_{n \in \mathbb{Z}}$ by

$$A^{(n)}(x) := A(x_{n-1}, x_n) \dots A(x_1, x_2) A(x_0, x_1).$$

The cocycle F_A is determined by the data (K, μ, A) , and identified by the function A , in contexts where the Markov system (K, μ) is fixed.

Definition 1.9. Let $\text{Gr}(\mathbb{R}^m)$ denote the Grassmann manifold of the Euclidean space \mathbb{R}^m . A \mathcal{F} -measurable section $V : \Sigma \rightarrow \text{Gr}(\mathbb{R}^m)$ is called A -invariant when

$$A(x_{n-1}, x_n) V(x_{n-1}) = V(x_n) \text{ for } \mathbb{P}_\mu\text{-a.e. } x = (x_n)_{n \in \mathbb{Z}}.$$

Assuming (K, μ) is strongly mixing, the ergodicity of this Markov kernel implies that the subspaces $V(x)$ have constant dimension μ -a.e., denoted by $\dim(V)$. We say that this family is *proper* if $0 < \dim(V) < d$.

Next we introduce the concepts of irreducible and totally irreducible cocycle (see definition 2.7 in [1]).

Definition 1.10. A cocycle $A \in \mathcal{B}_m^\infty(K)$ is called *irreducible w.r.t. (K, μ)* if it admits no measurable proper A -invariant section $V : \Sigma \rightarrow \text{Gr}(\mathbb{R}^m)$. A cocycle $A \in \mathcal{B}_m^\infty(K)$ is called *totally irreducible w.r.t. (K, μ)* if the exterior powers $\wedge_k A$ are irreducible for all $1 \leq k \leq m-1$.

We denote by $\mathcal{I}_m^\infty(K)$ the subspace of totally irreducible cocycles in $\mathcal{B}_m^\infty(K)$.

Proposition 1.3. The subspace $\mathcal{I}_m^\infty(K)$ is open in $\mathcal{B}_m^\infty(K)$.

Proof. A cocycle $A \in \mathcal{B}_m^\infty(K)$ is reducible (i.e. not irreducible) if it admits a measurable proper A -invariant section $V : \Sigma \rightarrow \text{Gr}(\mathbb{R}^m)$. It is enough to prove that the set of reducible cocycles is closed.

Let $A_k \rightarrow A$ be a convergent sequence of reducible cocycles in $\mathcal{B}_m^\infty(K)$, and let $V_k : \Sigma \rightarrow \text{Gr}(\mathbb{R}^m)$ be a measurable proper A_k -invariant section. We will prove that A is also reducible.

We will assume the probability space (Σ, μ) to be complete.

Let $\Omega \subset X$ be a Borel measurable set with $\mathbb{P}_\mu(\Omega) = 1$ such that for all $k \geq 1$ all $x = (x_n)_{n \in \mathbb{Z}} \in \Omega$ and $n \in \mathbb{Z}$, $A_k(x_{n-1}, x_n) V_k(x_{n-1}) = V_k(x_n)$.

Fix any point $s_0 \in \Sigma$. Extracting a subsequence we may assume that $V_k(s_0)$ converges to $V_0 \in \text{Gr}(\mathbb{R}^m)$ as k tends to ∞ . Consider then the set

$$A := \{s \in \Sigma : \exists x \in \Omega, n \in \mathbb{N} \text{ such that } x_0 = s_0 \text{ and } x_n = s\}.$$

In general A may fail to be a Borel set, but it is an *analytic set* in the sense of Descriptive set theory (see [12, Definition 14.1 and Exercise 14.3]). By [12, Theorem 21.10] this set is universally measurable, and in particular it is measurable w.r.t. μ . Hence, because of the strong mixing property,

$$\begin{aligned} \mu(\Sigma \setminus A) &= \lim_{n \rightarrow \infty} \mathbb{E}_{s_0}[\mathbf{1}_{\Sigma \setminus A}(e_n)] \\ &= \lim_{n \rightarrow \infty} \mathbb{P}_{s_0}\{x \in \Omega : e_n(x) \in \Sigma \setminus A\} = \lim_{n \rightarrow \infty} \mathbb{P}_{s_0}(\emptyset) = 0, \end{aligned}$$

which proves that for μ -a.e. $s \in \Sigma$ there exists a sequence $x \in \Omega$ such that $x_0 = s_0$ and $x_n = s$ for some $n \in \mathbb{N}$.

Then $V_k(s) = A_k(x_{n-1}, x_n) \dots A_k(x_1, x_2) A_k(x_0, x_1) V_k(s_0)$, which implies that $V_k(s)$ converges to $A(x_{n-1}, x_n) \dots A(x_1, x_2) A(x_0, x_1) V_0$ when $k \rightarrow \infty$. Thus, $V_k(s)$ converges for μ -a.e. $s \in \Sigma$, and the limit function $V(s) = \lim_{k \rightarrow \infty} V_k(s)$ is a measurable and proper A -invariant section, with the same dimension as the sections V_k . This proves that the cocycle A is reducible. \square

For the reader's convenience we briefly recall some definitions and notations regarding the Lyapunov exponents, Oseledets filtrations and decompositions of a cocycle A in any space of cocycles \mathcal{C}_m .

The ergodic theorem of Kingman allows us to define the Lyapunov exponents $L_j(A)$ with $1 \leq j \leq m$ as $L_j(A) := \Lambda_j(A) - \Lambda_{j-1}(A)$ where

$$\Lambda_j(A) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \|\wedge_j A(x)\| \quad \text{for } \mu\text{-a.e. } x \in X.$$

Let $\tau = (1 \leq \tau_1 < \dots < \tau_k < m)$ be a signature. If $A \in \mathcal{C}_m$ has a τ -gap pattern, i.e., $L_{\tau_j}(A) > L_{\tau_{j+1}}(A)$ for all j , we define the Lyapunov τ -block

$$\Lambda^\tau(A) := (\Lambda_{\tau_1}(A), \dots, \Lambda_{\tau_k}(A)) \in \mathbb{R}^k.$$

A flag of \mathbb{R}^m is any increasing sequence of linear subspaces. The corresponding sequence of dimensions is called its signature. A measurable filtration is a measurable function on X , taking values in the space of flags of \mathbb{R}^m with almost sure constant signature. We denote by $\mathfrak{F}(X, \mathbb{R}^m)$ the space of measurable filtrations. Note that the Oseledets filtration of A , which we denote by $F(A)$, is an element of this space.

We denote by $\mathfrak{F}_{\supset \tau}(X, \mathbb{R}^m)$ the subset of measurable filtrations with a signature τ or finer. If $F \in \mathfrak{F}_{\supset \tau}(X, \mathbb{R}^m)$ there is a natural projection F^τ with signature τ , obtained from F by simply ‘forgetting’ some of its components. This space is endowed with the following pseudo-metric

$$\text{dist}_\tau(F, F') := \int_X d_\tau(F^\tau(x), (F')^\tau(x)) \mu(dx),$$

where d_τ refers to the metric on the τ -flag manifold.

On the space $\mathfrak{F}(X, \mathbb{R}^m)$ we consider the coarsest topology that makes the sets $\mathfrak{F}_{\supset \tau}(X, \mathbb{R}^m)$ open, and the pseudo-metrics dist_τ continuous.

A decomposition of \mathbb{R}^m is a sequence of linear subspaces $\{E_j\}_{1 \leq j \leq k+1}$ whose direct sum is \mathbb{R}^m . This determines the flag $E_1 \subset E_1 \oplus E_2 \subset \dots \subset E_1 \oplus \dots \oplus E_k$, whose signature τ also designates the signature of the decomposition.

A measurable decomposition is a measurable function on X , taking values in the space of decompositions of \mathbb{R}^m with almost sure constant signature. We denote by $\mathfrak{D}(X, \mathbb{R}^m)$ the space of measurable decompositions. Note that the Oseledets decomposition of A , which we denote by $E(A)$, is an element of this space.

We denote by $\mathfrak{D}_{\supset \tau}(X, \mathbb{R}^m)$ the subset of measurable decompositions with a signature τ or finer. If $E \in \mathfrak{D}_{\supset \tau}(X, \mathbb{R}^m)$ there is a natural restriction E^τ with signature τ , obtained from E by simply ‘patching up’ the appropriate components. This space is endowed with the following pseudo-metric

$$\text{dist}_\tau(E, E') := \int_X d_\tau(E^\tau(x), (E')^\tau(x)) \mu(dx),$$

where d_τ refers to the metric on the manifold of τ -decompositions.

On the space $\mathfrak{D}(X, \mathbb{R}^m)$ we consider the coarsest topology that makes the sets $\mathfrak{D}_{\sup}(X, \mathbb{R}^m)$ open, and the pseudo-metrics dist_τ continuous.

We are ready to state a general result on the continuity of the LE, the Oseledets filtration and the Oseledets decomposition for irreducible Markov cocycles.

Theorem 1.1. *Let (K, μ) be a strongly mixing Markov system and let $m \geq 1$.*

Then all Lyapunov exponents $L_j : \mathcal{I}_m^\infty(K) \rightarrow \mathbb{R}$, with $1 \leq j \leq m$, the Oseledets filtration $F : \mathcal{I}_m^\infty(K) \rightarrow \mathfrak{F}(X, \mathbb{R}^m)$, and the Oseledets decomposition $E : \mathcal{I}_m^\infty(K) \rightarrow \mathfrak{D}(X, \mathbb{R}^m)$, are continuous functions of the cocycle $A \in \mathcal{I}_m^\infty(K)$.

Moreover, if $A \in \mathcal{I}_m^\infty(K)$ has a τ -gap pattern then the functions Λ^τ , F^τ and E^τ are Hölder continuous in a neighborhood of A .

This theorem is proved in section 4.1. It is an application of theorems 3.1, 4.7 and 4.8 in [7]. The main ingredients in these applications are two theorems on base and fiber uniform LDT estimates of exponential type that we now formulate.

We begin with the base LDT theorem. Consider the metric $\tilde{d} : X \times X \rightarrow [0, 1]$

$$\tilde{d}(x, x') := 2^{-\inf\{|k| : k \in \mathbb{Z}, x_k \neq x'_k\}} , \quad (1)$$

for all $x = (x_k)_{k \in \mathbb{Z}}$ and $x' = (x'_k)_{k \in \mathbb{Z}}$ in X . Notice that X is not compact for the topology induced by \tilde{d} , unless Σ is finite. Given $k \in \mathbb{N}$, $\alpha > 0$ and $f \in L^\infty(X)$ define

$$\begin{aligned} v_k(f) &:= \sup\{|f(x) - f(y)| : \tilde{d}(x, y) \leq 2^{-k}\} , \\ v_\alpha(f) &:= \sup\{2^{\alpha k} v_k(f) : k \in \mathbb{N}\} , \\ \|f\|_\alpha &:= \|f\|_\infty + v_\alpha(f) , \\ \mathcal{H}_\alpha(X) &:= \{f \in L^\infty(X) : v_\alpha(f) < +\infty\} . \end{aligned}$$

The last set, $\mathcal{H}_\alpha(X)$, is the space of Hölder continuous functions with exponent α w.r.t. the distance \tilde{d} on X . In fact it follows easily from the definition that

$$v_\alpha(f) = \sup_{x \neq x'} \frac{|f(x) - f(x')|}{\tilde{d}(x, x')^\alpha} .$$

Proposition 1.4. *For all $0 \leq \alpha \leq 1$, $(\mathcal{H}_\alpha(X), \|\cdot\|_\alpha)$ is a unital Banach algebra, and also a lattice.*

Proof. To see that $(\mathcal{H}_\alpha(X), \|\cdot\|_\alpha)$ is a normed algebra with unity, it is enough to verify the following inequalities:

$$\begin{aligned} v_k(fg) &\leq \|f\|_\infty v_k(g) + \|g\|_\infty v_k(f) , \\ v_\alpha(fg) &\leq \|f\|_\infty v_\alpha(g) + \|g\|_\infty v_\alpha(f) . \end{aligned}$$

They imply that

$$\|fg\|_\alpha \leq \|f\|_\alpha \|g\|_\alpha ,$$

and clearly $\|\mathbf{1}\|_\alpha = \|\mathbf{1}\|_\infty + v_\alpha(\mathbf{1}) = 1 + 0 = 1$. The proof that $(\mathcal{H}_\alpha(X), \|\cdot\|_\alpha)$ is a lattice and a Banach space is left as an exercise. \square

Definition 1.11. *We say that $f : X \rightarrow \mathbb{C}$ is future independent if $f(x) = f(y)$ for any $x, y \in X$ such that $x_k = y_k$ for all $k \leq 0$. Define the space*

$$\mathcal{H}_\alpha(X^-) := \{ f \in \mathcal{H}_\alpha(X) : f \text{ is future independent} \} . \quad (2)$$

The space $\mathcal{H}_\alpha(X^-)$ is a closed sub-algebra of $\mathcal{H}_\alpha(X)$, and hence a unital Banach algebra itself.

Denote by \mathcal{F}^+ the sub σ -field of \mathcal{F} generated by cylinders in non-negative coordinates. Likewise, denote by \mathcal{F}^- the sub σ -field of \mathcal{F} generated by cylinders in non-positive coordinates. With this terminology, the subspace $\mathcal{H}_\alpha(X^-)$ consists of all \mathcal{F}^- -measurable functions in $\mathcal{H}_\alpha(X)$.

The base LDT theorem below makes use of the standard notation $\mathbb{E}_\mu(\xi) = \int_X \xi d\mu$. This theorem is proved in section 3.1.

Theorem 1.2. *Let (K, μ) be a strongly mixing Markov system. For any $0 < \alpha \leq 1$ and $\xi \in \mathcal{H}_\alpha(X^-)$ there exist $C = C(\xi) > 0$, $k = k(\xi) > 0$ and $\varepsilon_0 = \varepsilon_0(\xi) > 0$ such that for all $0 < \varepsilon < \varepsilon_0$, $x \in \Sigma$ and $n \in \mathbb{N}$,*

$$\mathbb{P}_\mu \left[\left| \frac{1}{n} \sum_{j=0}^{n-1} \xi \circ T^j - \mathbb{E}_\mu(\xi) \right| > \varepsilon \right] \leq C e^{-k\varepsilon^2 n} .$$

Moreover, the constants C , k and ε_0 depend only on K and $\|\xi\|_\alpha$, and hence can be kept constant when K is fixed and ξ ranges over any bounded set in $\mathcal{H}_\alpha(X^-)$.

The fiber LDT theorem, proved in section 3.2, has the following statement.

Theorem 1.3. *Given a Markov system (K, μ) and $A \in \mathcal{B}_m^\infty(K)$, assume*

- (1) (K, μ) is strongly mixing,
- (2) A is irreducible,
- (3) $L_1(A) > L_2(A)$.

Then there exists \mathcal{V} neighborhood of A in $\mathcal{B}_m^\infty(K)$ and there exist $C > 0$, $k > 0$ and $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$, $B \in \mathcal{V}$ and $n \in \mathbb{N}$,

$$\mathbb{P}_\mu \left[\left| \frac{1}{n} \log \|B^{(n)}\| - L_1(B) \right| > \varepsilon \right] \leq C e^{-k\varepsilon^2 n}.$$

1.2. The spectral method. Consider a Markov system (K, μ) on a compact metric space Σ . Given some \mathcal{F} -measurable observable $\xi : \Sigma \rightarrow \mathbb{R}$, let $\hat{\xi} : X^+ \rightarrow \mathbb{R}$ be the \mathcal{F}^+ -measurable function $\hat{\xi}(x) = \xi(x_0)$.

Given $x \in \Sigma$, let \mathbb{P}_x^+ denote the probability on the measurable space (X^+, \mathcal{F}^+) that makes $\{e_n : X^+ \rightarrow \Sigma\}_{n \geq 0}$ a Markov process with transition kernel K and initial distribution with point mass δ_x (see Definition 1.3). Then $\{\hat{\xi} \circ T^n\}_{n \geq 0}$ is also a Markov process on $(X^+, \mathcal{F}^+, \mathbb{P}_x^+)$.

Definition 1.12. We call sum process to the following sequence of random variables $\{S_n(\xi)\}_{n \geq 0}$ on (X^+, \mathcal{F}^+) ,

$$S_n(\xi)(x) := \sum_{j=0}^{n-1} \hat{\xi} \circ T^j(x) = \sum_{j=0}^{n-1} \xi(x_j).$$

Definition 1.13. An observed Markov system on (Σ, \mathcal{F}) is a triple (K, μ, ξ) where (K, μ) is a Markov system on (Σ, \mathcal{F}) , and $\xi : \Sigma \rightarrow \mathbb{R}$ is an \mathcal{F} -measurable function.

Definition 1.14. We say that ξ satisfies LDT estimates of exponential type if there exist positive constants C , k and ε_0 such that for all $n \in \mathbb{N}$, $0 < \varepsilon < \varepsilon_0$ and $x \in \Sigma$,

$$\mathbb{P}_x^+ \left\{ y \in X^+ : \left| \frac{1}{n} S_n(\xi)(y) - \mathbb{E}_\mu(\xi) \right| > \varepsilon \right\} \leq C e^{-nk\varepsilon^2}.$$

Given a class \mathfrak{X} of observed Markov systems (K, μ, ξ) on a given measurable space (Σ, \mathcal{F}) , we say that \mathfrak{X} satisfies uniform LDT estimates of exponential type if there exist positive constants C , k and ε_0 such that for every observed Markov system $(K, \mu, \xi) \in \mathfrak{X}$, the observable ξ satisfies LDT estimates of exponential type with constants C , k and ε_0 .

Definition 1.15. Let $\eta : X^+ \rightarrow \mathbb{R}$ be a random variable on (X^+, \mathcal{F}^+) . The function $c(\eta, x, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$,

$$c(\eta, x, t) := \log \mathbb{E}_x[e^{t\eta}]$$

is called the second characteristic function of η , also known as the cumulant generating function of η (see [15]).

Proposition 1.5. *Let $\eta : X^+ \rightarrow \mathbb{R}$ be a \mathcal{F}^+ -measurable random variable. Assume there exist $a, M > 0$ such that for all $x \in \Sigma$, $\max\{\mathbb{E}_x[e^{a\eta}], \mathbb{E}_x[e^{a\eta}|\eta|\}] \leq M$. Then the cumulant generating function $c(\eta, x, \cdot)$ satisfies*

- (1) $c(\eta, x, t)$ is well-defined and analytic for $t \in (-a, a)$,
- (2) $c(\eta, x, 0) = 0$,
- (3) $\frac{dc}{dt}(\eta, x, 0) = \mathbb{E}_x(\eta)$,
- (4) $c(\eta, x, t) \geq t \mathbb{E}_x(\eta)$, for all $t \in (-a, a)$,
- (5) the function $c(\eta, x, \cdot) : (-a, a) \rightarrow \mathbb{R}$, $t \mapsto c(\eta, x, t)$, is convex.

Proof. For (1) notice that the assumptions imply that the parametric integral $\mathbb{E}_x(e^{z\eta})$ and its formal derivative $\mathbb{E}_x(e^{z\eta}\eta)$ are well-defined continuous functions on the disk $|z| < a$. Since $c(\eta, x, 0) = \log \mathbb{E}_x(\mathbf{1}) = \log 1 = 0$, (2) follows. Property (3) holds because $\frac{dc}{dt}(\eta, x, 0) = \mathbb{E}_x(\eta \mathbf{1}) / \mathbb{E}_x(\mathbf{1}) = \mathbb{E}_x(\eta)$. The convexity (5) follows by Hölder inequality, with conjugate exponents $p = 1/s$ and $q = 1/(1-s)$, where $0 < s < 1$. In fact, for all $t_1, t_2 \in \mathbb{R}$,

$$\begin{aligned} c(\eta, x, s t_1 + (1-s) t_2) &= \log \mathbb{E}_x[(e^{t_1 \eta})^s (e^{t_2 \eta})^{1-s}] \\ &\leq \log (\mathbb{E}_x[e^{t_1 \eta}]^s (\mathbb{E}_x[e^{t_2 \eta}])^{1-s}) \\ &= s c(\eta, x, t_1) + (1-s) c(\eta, x, t_2). \end{aligned}$$

Finally, (2), (3) and (5) imply (4). \square

Given an observable $\xi : \Sigma \rightarrow \mathbb{R}$, the function $c_n(\xi, x, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$c_n(\xi, x, t) := \log \mathbb{E}_x[e^{t S_n(\xi)}],$$

is the cumulant generating function of $S_n(\xi)$. Under general conditions, e.g., if ξ is bounded, this function is analytic in \mathbb{C} , or at least analytic in a neighbourhood of 0.

Let us write $\mathbb{D}_a(0) = \{z \in \mathbb{C} : |z| < a\}$.

Definition 1.16. *We call limit cumulant generating function of the process $\{S_n(\xi)\}_{n \geq 0}$ to any function $c(\xi, \cdot) : \mathbb{D}_a(0) \rightarrow \mathbb{C}$ such that there exist a constant $C > 0$ and a numeric sequence $\{\delta_n\}_{n \geq 0}$ for which the following properties hold:*

- (1) $c_n(\xi, \cdot)$ is well defined and analytic on $\mathbb{D}_a(0)$, for all $n \in \mathbb{N}$,
- (2) $|n c(\xi, z) - c_n(\xi, x, z)| \leq C |z| + \delta_n$, for all $n \in \mathbb{N}$, $z \in \mathbb{D}_a(0)$ and $x \in \Sigma$,
- (3) $\lim_{n \rightarrow +\infty} \delta_n = 0$.

Before discussing why they exist, let us draw some conclusions from the existence of limit cumulant generating functions.

Proposition 1.6. *Given an \mathcal{F} -measurable observable $\xi : \Sigma \rightarrow \mathbb{R}$, let $c(\xi, z)$ be a limit cumulant generating function of the process $\{S_n(\xi)\}_{n \geq 0}$ on $\mathbb{D}_a(0)$. Then*

- (1) $z \mapsto c(\xi, z)$ is analytic on $\mathbb{D}_a(0)$,
- (2) $c(\xi, 0) = 0$,
- (3) $\frac{dc}{dt}(\xi, 0) = \mathbb{E}_\mu(\hat{\xi})$,
- (4) $c(\xi, t) \geq t \mathbb{E}_\mu(\hat{\xi})$, for all $t \in \mathbb{R}$,
- (5) the function $c(\xi, \cdot) : (-a, a) \rightarrow \mathbb{R}$, $t \mapsto c(\xi, t)$, is convex.

Proof. The function $c(\xi, z)$ is analytic on $\mathbb{D}_a(0)$ because it is the uniform limit of the sequence of analytic functions $\frac{1}{n} c_n(\xi, x, z)$. This proves (1).

Item (2) follows directly from proposition 1.5 (2).

Consider now the sequence of analytic functions

$$\hat{c}_n(\xi, z) := \int_{\Sigma} c_n(\xi, x, z) d\mu(x) .$$

Then

$$\frac{d\hat{c}_n}{dt}(\xi, 0) = \int_{\Sigma} \mathbb{E}_x[S_n(\xi)] d\mu(x) = \mathbb{E}_\mu(\hat{\xi}) .$$

Taking the limit identity (3) holds.

Since convexity is a closed property, (5) follows from proposition 1.5 (5).

Finally, (2),(3) and (5) imply (4). \square

Next proposition relates the existence of a limit cumulant generating function for the process $\{S_n(\xi)\}_{n \geq 0}$ with LDT estimates of exponential type for ξ .

Proposition 1.7. *Let $\xi : \Sigma \rightarrow \mathbb{R}$ be \mathcal{F} -measurable observable, and $c(\xi, z)$ be a limit cumulant generating function of the process $\{S_n(\xi)\}_{n \geq 0}$ on $\mathbb{D}_a(0)$.*

Given $h > \frac{d^2 c}{dt^2}(\xi, 0)$, there exist $C, \varepsilon_0 > 0$ such that for all $n \in \mathbb{N}$, $x \in \Sigma$ and $0 < \varepsilon < \varepsilon_0$,

$$\mathbb{P}_x^+ \left[\left| \frac{1}{n} S_n(\xi) - \mathbb{E}_\mu(\hat{\xi}) \right| > \varepsilon \right] \leq C e^{-n \frac{\varepsilon^2}{2h}} .$$

In other words, ξ satisfies LDT estimates of exponential type.

Proof. Let us abbreviate $c(t) = c(\xi, t)$. We can assume that $c'(0) = \mathbb{E}_\mu(\hat{\xi}) = 0$. Otherwise we would work with $\xi' = \xi - \mathbb{E}_\mu(\xi) \mathbf{1}$, for which $\mathbb{E}_\mu(\xi') = 0$. Notice that the normalized process $\{S_n(\xi')\}_{n \geq 0}$ admits the limit cumulant generating function $c(\xi', t) = c(t) - t \mathbb{E}_\mu(\xi) = c(t) - t c'(0)$.

Since $h > c''(0)$, we can choose $0 < t_0 < a$ such that for all $t \in (-t_0, t_0)$,

$$0 \leq c(t) < \frac{h t^2}{2}.$$

By definition 1.16, for all $t \in (-t_0, t_0)$,

$$\mathbb{E}_x[e^{t S_n(\xi)}] = e^{c_n(\xi, x, t)} \leq e^{n c(t) + C|t| + \delta_n} \leq \frac{C_0}{2} e^{n \frac{h t^2}{2}},$$

where $C_0 := 2 e^{C t_0 + \sup_{n \geq 0} \delta_n}$. Thus, by Chebyshev's inequality we have for all $|t| < t_0$

$$\mathbb{P}_x[S_n(\xi) > n\varepsilon] \leq e^{-tn\varepsilon} \mathbb{E}_x[e^{t S_n(\xi)}] \leq \frac{C_0}{2} e^{-n \left(t\varepsilon - \frac{h t^2}{2} \right)}.$$

Given $0 < \varepsilon < \varepsilon_0 := h t_0$, pick $t = \frac{\varepsilon}{h} \in]0, t_0[$. This choice of t minimizes the function $g(t) = e^{-\left(t\varepsilon - \frac{h t^2}{2} \right)}$. For this value of t we obtain

$$\mathbb{P}_x[S_n(\xi) > n\varepsilon] \leq \frac{C_0}{2} e^{-\frac{\varepsilon^2}{2h} n}.$$

We can derive the same conclusion for $-\xi$, because $c(\xi, -t)$ is a limit cumulant generating function of the process $\{S_n(-\xi)\}_{n \geq 0}$,

$$\mathbb{P}_x[S_n(\xi) < -n\varepsilon] = \mathbb{P}_x[S_n(-\xi) > n\varepsilon] \leq \frac{1}{2} C_0 e^{-\frac{\varepsilon^2}{2h} n}.$$

Thus, for all $x \in \Sigma$, $0 < \varepsilon < \varepsilon_0$ and $n \in \mathbb{N}$,

$$\mathbb{P}_x[|S_n(\xi)| > n\varepsilon] \leq C_0 e^{-\frac{\varepsilon^2}{2h} n}.$$

□

Remark 1.1. *To obtain a sharp upper bound on the rate function for the large deviations of the process $S_n(\xi)$ we should have used the Legendre transform of the convex function $c(t) - t c'(0)$. Here because we do not care about sharp estimates, but mainly to avoid dealing with the degenerate case where $c(t)$ is not strictly convex, we have replaced $c(t) - t c'(0)$ by its upper bound $\frac{h t^2}{2}$ on the small neighborhood $(-t_0, t_0)$, which is always strictly convex.*

Consider now a topological space \mathfrak{X} of observed Markov systems (K, μ, ξ) , on a given measurable space (Σ, \mathcal{F}) .

Denote by $\mathcal{H}(\mathbb{D}_a(0))$ the Banach space of analytic functions $f : \mathbb{D}_a(0) \rightarrow \mathbb{C}$ with a continuous extension up to the disk's closure. Endow this space with the usual max norm $\|f\|_\infty = \max_{|z| \leq a} |f(z)|$.

Corollary 1.8. *Assume there is continuous map $c : \mathfrak{X} \rightarrow \mathcal{H}(\mathbb{D}_a(0))$ such that*

- (a) for each $(K, \mu, \xi) \in \mathfrak{X}$, the function $c(\xi, z) := c(K, \mu, \xi)(z)$ is a limit cumulant generating function of the process $\{S_n(\xi)\}_{n \geq 0}$ on $\mathbb{D}_a(0)$,
- (b) the parameters C and δ_n in definition 1.16 can be chosen uniformly in \mathfrak{X} .

Then

- (1) For each $(K, \mu, \xi) \in \mathfrak{X}$ there exists a neighborhood \mathcal{V} in \mathfrak{X} such that \mathcal{V} satisfies uniform LDT estimates of exponential type.
- (2) If there exists $h > 0$ such that $\frac{d^2}{dt^2} c(\xi, 0) < h$ for all $(K, \mu, \xi) \in \mathfrak{X}$ then \mathfrak{X} satisfies uniform LDT estimates of exponential type.

Proof. Given $(K_0, \mu_0, \xi_0) \in \mathfrak{X}$, let $c_0(t) := c(K_0, \mu_0, \xi_0)(t)$, and take $h > c_0''(0)$. By continuity of $c : \mathfrak{X} \rightarrow \mathcal{H}(\mathbb{D}_a(0))$ there exist a neighborhood \mathcal{V} of (K_0, μ_0, ξ_0) in \mathfrak{X} and $t_0 > 0$ such that for any $(K, \mu, \xi) \in \mathcal{V}$, the function $c(\xi, z) := c(K, \mu, \xi)(z)$ satisfies for all $t \in (-t_0, t_0)$,

$$c(\xi, t) - t \frac{dc}{dt}(\xi, 0) < \frac{h t^2}{2} .$$

The argument used to prove proposition 1.7 shows that \mathcal{V} satisfies uniform LDT estimates of exponential type. \square

The strategy to meet the assumptions of corollary 1.8, i.e., to prove the existence of a limit cumulant generating function for the process $\{S_n(\xi)\}_{n \geq 0}$, is a spectral method that we describe now.

Define a family of Laplace-Markov operators

$$(Q_t f)(x) = (Q_{K, \xi, t} f)(x) := \int_{\Sigma} f(y) e^{t \xi(y)} K(x, dy) ,$$

on some appropriate Banach space \mathcal{B} , embedded in $L^\infty(\Sigma, \mathcal{F})$, and containing the constant functions. Notice that by definition $(Q_t \mathbf{1})(x) = \mathbb{E}_x[e^{t \hat{\xi}}]$. Hence, iterating this relation we obtain the following formula for the moment generating function of $S_n(\xi)$: for all $x \in \Sigma$ and $n \in \mathbb{N}$,

$$\mathbb{E}_x[e^{t S_n(\xi)}] = (Q_t^n \mathbf{1})(x) .$$

For $t = 0$, the operator $Q_0 : \mathcal{B} \rightarrow \mathcal{B}$, is a Markov operator. In particular it is a positive operator which fixes the constant functions, e.g., $Q_0 \mathbf{1} = \mathbf{1}$, and whose spectrum is contained in the closed unit disk. The key ingredient to estimate the moment generating function $\mathbb{E}_x[e^{t S_n(\xi)}]$ via this spectral approach is the assumption that the operator $Q_0 : \mathcal{B} \rightarrow \mathcal{B}$ is *quasi-compact* and *simple*. This means that the eigenvalue 1 of Q_0 is simple and there exists a spectral gap separating this eigenvalue from the rest of spectrum inside the open unit disk. Under this hypothesis, Q_t is a positive operator, whenever defined, and there exists a unique

eigenfunction $v(t) \in \mathcal{B}$ such that $Q_t v(t) = \lambda(t) v(t)$, normalized by $\mathbb{E}_\mu[v(t)] = 1$, and corresponding to a positive eigenvalue $\lambda(t)$ of Q_t . Hence, because the functions $t \mapsto \lambda(t)$ and $t \mapsto v(t)$ are continuous in t (in fact analytic), we have

$$\mathbb{E}_\mu[e^{t S_n(\xi)}] = \int (Q_t^n \mathbf{1}) d\mu \approx \int Q_t^n v(t) d\mu = \int \lambda(t)^n v(t) d\mu = \lambda(t)^n.$$

From this relation we infer that $c(t) = \log \lambda(t)$ is a limit cumulant generating function for the process $S_n(\xi)$. Therefore, by proposition 1.7, ξ satisfies LDT estimates of exponential type.

To obtain uniform LDT estimates, through corollary 1.8, we assume some weak continuous dependence of the family of operators $t \mapsto Q_{K,\xi,t}$ on the observed Markov system (K, μ, ξ) , which implies that the eigenvalue function $\lambda(t) \in \mathcal{H}(\mathbb{D}_a(0))$ also depends continuously on (K, μ, ξ) .

1.3. Literature review. We mention briefly some of the origins of this subject.

One is the aforementioned Furstenberg's work, started with the proof by H. Furstenberg and H. Kesten of a law of large numbers for random i.i.d. products of matrices [8], and later abstracted by Furstenberg to a seminal theory on random products in semisimple Lie groups [9]. In this context, a first central limit theorem was proved by V. N. Tutubalin in [20]. Since its origin, the scope of Furstenberg's theory has been greatly extended by many contributions. See for instance the book of A. Raugi [18] and Y. Guivarc'h and A. Raugi's paper [10].

Another source is a central limit theorem of S.V.Nagaev for stationary Markov chains (see [16]). In his approach Nagaev uses the spectral properties of a quasi-compact Markov operator acting on some space of bounded measurable functions. This method was used by E. Le Page to obtain more general central limit theorems, as well as a large deviation principle, for random i.i.d. products of matrices [14]. Later P. Bougerol extended Le Page's approach, proving similar results for Markov type random products of matrices (see [1]).

The book of P. Bougerol and J. Lacroix [2], on random i.i.d. products of matrices, is an excellent introduction on the subject in [14, 1]. More recently, the book of H. Hennion and L. Hervé [11] describes a powerful abstract setting where the method of Nagaev can be applied to derive limit theorems. It contains several applications, including to dynamical systems and linear cocycles, that illustrate the method.

2. AN ABSTRACT SETTING

In this section we specialize an abstract setting in [11], from which we derive an abstract theorem on the existence of uniform LDT estimates for Markov processes.

2.1. The assumptions. Let \mathcal{B} be a Banach space, and $\mathcal{L}(\mathcal{B})$ denote the Banach algebra of bounded linear operators $T : \mathcal{B} \rightarrow \mathcal{B}$. Given $T \in \mathcal{L}(\mathcal{B})$, we denote its spectrum by $\sigma(T)$, and its spectral radius by

$$\rho(T) = \lim_{n \rightarrow +\infty} \|T^n\|^{1/n} = \inf_{n \geq 0} \|T^n\|^{1/n} .$$

Definition 2.1. *The operator T is called quasi-compact if there is a T -invariant decomposition $\mathcal{B} = F \oplus \mathcal{H}$ such that $\dim F < +\infty$ and the spectral radius of $T|_{\mathcal{H}}$ is (strictly) less than the absolute value $|\lambda|$ of any eigenvalue λ of $T|_F$. T is called quasi-compact and simple when furthermore $\dim F = 1$. In this case $\sigma(T|_F)$ consists of a single simple eigenvalue referred to as the maximal eigenvalue of T .*

Consider a Markov system (K, μ) on a compact metric space Σ .

Definition 2.2. *The following linear operator is called a Markov operator*

$$(Qf)(x) = (Q_K f)(x) := \int_X f(y) K(x, dy) .$$

It operates on \mathcal{F} -measurable functions on Σ , mapping L^p functions to L^p functions, for any $1 \leq p \leq \infty$. We shall write Q instead of Q_K when the kernel K is fixed.

Definition 2.3. *The following linear operator is called a Laplace-Markov operator*

$$(Q_\xi f)(x) = (Q_{K, \xi} f)(x) := \int_X f(y) e^{\xi(y)} K(x, dy) .$$

It also operates on \mathcal{F} -measurable functions on Σ , but the domain of Q_ξ depends also on the observable ξ .

Proposition 2.1. *Given a Markov system (K, μ) the following are equivalent:*

- (a) (K, μ) is strongly mixing,
- (b) $Q_K : L^\infty(\Sigma, \mathcal{F}) \rightarrow L^\infty(\Sigma, \mathcal{F})$ is quasi-compact and simple.

Proof. If (K, μ) is strongly mixing, by definition 1.6 there exist constants $C > 0$ and $0 < \rho < 1$ such that for all $f \in L^\infty(\Sigma)$,

$$\|(Q_K)^n f - \langle f, \mu \rangle \mathbf{1}\|_\infty \leq C \rho^n \|f\|_\infty .$$

Defining

$$H_0 = \{ f \in L^\infty(\Sigma) : \langle f, \mu \rangle = 0 \} ,$$

since $(Q_K)^* \mu = \mu$, this subspace is Q_K -invariant. Thus, we have a Q_K -invariant decomposition $L^\infty(\Sigma) = \langle \mathbf{1} \rangle \oplus H_0$ such that $\|(Q_K)^n|_{H_0}\| \leq C \rho^n$. This implies that $r(Q_K|_{H_0}) \leq \rho < 1$.

Conversely, if $Q_K : L^\infty(\Sigma) \rightarrow L^\infty(\Sigma)$ is quasi-compact and simple, there exists a Q_K -invariant decomposition $L^\infty(\Sigma) = \langle \mathbf{1} \rangle \oplus H_0$ such that $r(Q_K|_{H_0}) < 1$. By the Hahn-Banach Theorem there is a bounded linear functional $\Lambda : L^\infty(\Sigma) \rightarrow \mathbb{R}$ such that $\Lambda(\mathbf{1}) = 1$, and $\Lambda(f) = 0$ for all $f \in H_0$. We claim that Λ is positive functional, i.e., $\Lambda(f) \geq 0$ whenever $f \geq 0$. Take any function $f \in L^\infty(\Sigma)$ such that $f \geq 0$, and write $f = c \mathbf{1} + h$ with $h \in H_0$. Since Q_K is a positive operator we have

$$c \mathbf{1} = \lim_{n \rightarrow +\infty} (c \mathbf{1} + (Q_K)^n h) = \lim_{n \rightarrow +\infty} (Q_K)^n f \geq 0 ,$$

which implies that $c = \Lambda(f) \geq 0$. Hence Λ is positive. By the Riez-Markov-Kakutani Theorem there is a probability measure μ on Σ such that $\Lambda(f) = \int_\Sigma f d\mu$, for all $f \in L^\infty(\Sigma)$.

Let us prove that μ is K -stationary. Given $f \in L^\infty(\Sigma)$, write $f = c \mathbf{1} + h$, with $h \in H_0$. Hence $Q_K f = c \mathbf{1} + Q_K h$ with $Q_K h \in H_0$. This proves that μ is stationary,

$$\int_\Sigma (Q_K f) d\mu = \Lambda(Q_K f) = c = \Lambda(f) = \int_\Sigma f d\mu .$$

Now, because H_0 is the kernel of $\Lambda : L^\infty(\Sigma) \rightarrow \mathbb{R}$, we get that for all $f \in L^\infty(\Sigma)$, $f \in H_0 \Leftrightarrow \langle f, \mu \rangle = 0$. Thus $f - \langle f, \mu \rangle \mathbf{1} \in H_0$, and taking $r(Q_K|_{H_0}) < \rho < 1$, there is a constant $C > 0$ such that

$$\begin{aligned} \|(Q_K)^n f - \langle f, \mu \rangle \mathbf{1}\|_\infty &= \|(Q_K)^n [f - \langle f, \mu \rangle \mathbf{1}]\|_\infty \\ &\leq C \rho^n \|f - \langle f, \mu \rangle \mathbf{1}\|_\infty \\ &\leq 2 C \rho^n \|f\|_\infty . \end{aligned}$$

This proves that (K, μ) is strongly mixing. \square

We discuss now a setting, consisting of the assumptions (B1)-(B7) and (A1)-(A4) below, where an abstract LDT theorem is proved, and from which theorems 1.2 and 1.3 will be deduced. The context here specializes a more general setting in [11].

Let $(\mathfrak{X}, \text{dist})$ be a metric space of observed Markov systems (K, μ, ξ) over the compact metric space (Σ, d) . Besides \mathfrak{X} , this setting consists of a scale of complex Banach algebras $(\mathcal{B}_\alpha, \|\cdot\|_\alpha)$ indexed in $\alpha \in [0, 1]$, where each \mathcal{B}_α is a space of bounded Borel measurable functions on Σ . We assume that there exist seminorms $v_\alpha : \mathcal{B}_\alpha \rightarrow [0, +\infty[$ such that for all $0 \leq \alpha \leq 1$,

- (B1) $\|f\|_\alpha = v_\alpha(f) + \|f\|_\infty$, for all $f \in \mathcal{B}_\alpha$,
- (B2) $\mathcal{B}_0 = L^\infty(\Sigma)$, and $\|\cdot\|_0$ is equivalent to $\|\cdot\|_\infty$,
- (B3) \mathcal{B}_α is a lattice, i.e., if $f \in \mathcal{B}_\alpha$ then $\overline{f}, |f| \in \mathcal{B}_\alpha$,
- (B4) \mathcal{B}_α is a Banach algebra with unity $\mathbf{1} \in \mathcal{B}_\alpha$ and $v_\alpha(\mathbf{1}) = 0$.

Assume also that this family is a *scale of normed spaces* in the sense that for all $0 \leq \alpha_0 < \alpha_1 < \alpha_2 \leq 1$ (see [13])

- (B5) $\mathcal{B}_{\alpha_2} \subset \mathcal{B}_{\alpha_1} \subset \mathcal{B}_{\alpha_0}$,
- (B6) $v_{\alpha_0}(f) \leq v_{\alpha_1}(f) \leq v_{\alpha_2}(f)$, for all $f \in \mathcal{B}_{\alpha_2}$,
- (B7) $v_{\alpha_1}(f) \leq v_{\alpha_0}(f)^{\frac{\alpha_2 - \alpha_1}{\alpha_2 - \alpha_0}} v_{\alpha_2}(f)^{\frac{\alpha_1 - \alpha_0}{\alpha_2 - \alpha_0}}$, for all $f \in \mathcal{B}_{\alpha_2}$.

An example of a scale of Banach algebras satisfying (B1)-(B7) are the spaces of α -Hölder continuous functions on (Σ, d) . The norms on these spaces are defined as follows: for all $\alpha \in]0, 1]$ and $f \in L^\infty(\Sigma)$, let

$$\|f\|_\alpha := v_\alpha(f) + \|f\|_\infty, \quad \text{with} \quad v_\alpha(f) := \sup_{\substack{x, y \in \Sigma \\ x \neq y}} \frac{|f(x) - f(y)|}{d(x, y)^\alpha}.$$

Proposition 2.2. *If (Σ, d) has diameter ≤ 1 then the family of spaces*

$$\mathcal{H}_\alpha(\Sigma) := \{f \in L^\infty(\Sigma) : v_\alpha(f) < +\infty\}, \quad \alpha \in [0, 1]$$

satisfies (B1)-(B7).

Proof. (B1) holds by definition of the Hölder norm $\|\cdot\|_\alpha$. For (B2) notice that $v_0(f)$ measures the oscillation of f , and hence $v_0(f) \leq 2\|f\|_\infty$. Property (B3) is obvious. Assumption (B4) follows from the following inequality

$$v_\alpha(fg) \leq \|f\|_\infty v_\alpha(g) + \|g\|_\infty v_\alpha(f),$$

that holds for all $f, g \in L^\infty(\Sigma)$. The monotonicity properties (B5) and (B6) are straightforward to check. Finally, assumption (B7) follows from the convexity of the function $\alpha \mapsto \log v_\alpha(f)$. Given $\alpha_1, \alpha_2, s \in [0, 1]$,

$$\begin{aligned} \log v_{s\alpha_1 + (1-s)\alpha_2}(f) &= \log \sup_{x \neq y} \frac{|f(x) - f(y)|^{s+(1-s)}}{d(x, y)^{s\alpha_1 + (1-s)\alpha_2}} \\ &\leq \log \left(\sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)^{\alpha_1}} \right)^s \left(\sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)^{\alpha_2}} \right)^{1-s} \\ &= s \log v_{\alpha_1}(f) + (1-s) \log v_{\alpha_2}(f). \end{aligned}$$

□

We make now a second set of assumptions that rule the action of the Markov operators, associated to observable Markov systems $(K, \mu, \xi) \in \mathfrak{X}$, on the Banach algebras \mathcal{B}_α .

Assume there exists an interval $[\alpha_1, \alpha_0] \subset (0, 1]$ with $\alpha_1 < \frac{\alpha_0}{2}$ such that for all $\alpha \in [\alpha_1, \alpha_0]$ the following properties hold:

- (A1) $(K, \mu, -\xi) \in \mathfrak{X}$ whenever $(K, \mu, \xi) \in \mathfrak{X}$.
- (A2) The Markov operators $Q_K : \mathcal{B}_\alpha \rightarrow \mathcal{B}_\alpha$ are uniformly quasi-compact and simple. More precisely, there exist constants $C > 0$ and $0 < \sigma < 1$ such that for all $(K, \mu, \xi) \in \mathfrak{X}$ and $f \in \mathcal{B}_\alpha$,

$$\|Q_K^n f - \langle f, \mu \rangle \mathbf{1}\|_\alpha \leq C \sigma^n \|f\|_\alpha.$$

- (A3) The operators $Q_{K,z\xi}$ act continuously on the Banach algebras \mathcal{B}_α , uniformly in $(K, \mu, \xi) \in \mathfrak{X}$ and z small. More precisely, we assume there are constants $b > 0$ and $M > 0$ such that for $i = 0, 1, 2$, $|z| < b$ and $f \in \mathcal{B}_\alpha$,

$$Q_{K,z\xi}(f \xi^i) \in \mathcal{B}_\alpha \quad \text{and} \quad \|Q_{K,z\xi}(f \xi^i)\|_\alpha \leq M \|f\|_\alpha.$$

- (A4) The family of functions $\mathfrak{X} \ni (K, \mu, \xi) \mapsto Q_{K,z\xi}$, indexed in $|z| \leq b$, is Hölder equi-continuous in the sense that there exists $0 < \theta \leq 1$ such that for all $|z| \leq b$, $f \in \mathcal{B}_\alpha$ and $(K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2) \in \mathfrak{X}$,

$$\|Q_{K_1,z\xi_1} f - Q_{K_2,z\xi_2} f\|_\infty \leq M \|f\|_\alpha \text{dist}((K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2))^\theta.$$

The interval $[\alpha_1, \alpha_0]$ will be called as the *range* of the scale of Banach algebras. In the fiber LDT theorem we will need to take α_0 small enough to have contraction in (A2), but at the same time we need α_1 bounded away from 0 to have uniformity in this contraction. The need for the condition $\alpha_1 < \frac{\alpha_0}{2}$ is explained in remark 2.1.

The positive constants C , σ , M , b and θ above will be called the *setting constants*.

Examples of contexts satisfying all assumptions (B1)-(B7) and (A1)-(A4) are provided by the applications in sections 3.1 and 3.2.

The symmetry assumption (A1) allows us to reduce deviations below average to deviations above average, thus shortening the arguments.

(A2) is the main assumption: all Markov operators $Q_K : \mathcal{B}_\alpha \rightarrow \mathcal{B}_\alpha$ are quasi-compact and simple, uniformly in $(K, \mu, \xi) \in \mathfrak{X}$. This will imply that, possibly decreasing b , all Laplace-Markov operators $Q_{K,z\xi} : \mathcal{B}_\alpha \rightarrow \mathcal{B}_\alpha$ are also quasi-compact and simple, uniformly in $(K, \mu, \xi) \in \mathfrak{X}$ and $|z| < b$.

(A3) is a regularity assumption. The operators $Q_{K,z\xi}$ act continuously on \mathcal{B}_α , uniformly in $(K, \mu, \xi) \in \mathfrak{X}$ and $|z| < b$. Moreover, it implies that $\mathbb{D}_b \ni z \mapsto Q_{K,z\xi} \in \mathcal{L}(\mathcal{B}_\alpha)$, is an analytic function.

Finally, (A4) implies that the function $(K, \mu, \xi) \mapsto \lambda_{K,\xi}(z)$ is uniformly Hölder continuous. Here $\lambda_{K,\xi}(z)$ denotes the maximal eigenvalue of $Q_{K,z\xi}$,

These facts follow from the propositions stated and proved in the rest of this subsection.

Hypothesis (A3) implies that $Q_{K,z\xi} \in \mathcal{L}(\mathcal{B}_\alpha)$, for all $z \in \mathbb{D}_b$. In particular the function $Q_{K,*\xi} : \mathbb{D}_b \rightarrow \mathcal{L}(\mathcal{B}_\alpha)$, $z \mapsto Q_{K,z\xi}$, is well-defined, for every $(K, \mu, \xi) \in \mathfrak{X}$.

Proposition 2.3. *The function $Q_{K,*\xi} : \mathbb{D}_b \rightarrow \mathcal{L}(\mathcal{B}_\alpha)$ is analytic with*

$$\frac{d}{dz} Q_{K,z\xi}(f) = Q_{K,z\xi}(f\xi) \text{ for } f \in \mathcal{B}_\alpha,$$

for all $(K, \mu, \xi) \in \mathfrak{X}$, and $\alpha_1 \leq \alpha \leq \alpha_0$.

Proof. Given $b \in \mathbb{R}$, for all $z, z_0 \in \mathbb{C}$,

$$\frac{e^{zb} - e^{z_0b}}{z - z_0} - b e^{z_0b} = \int_{z_0}^z b^2 e^{\zeta b} \frac{z - \zeta}{z - z_0} d\zeta.$$

This is the first order Taylor remainder formula for $h(z) = e^{bz}$ at $z = z_0$. To shorten notation we write Q_z for $Q_{K,z\xi}$. Replacing b by $\xi(y)$, multiplying by $f(y) K(x, dy)$ and integrating over Σ we get

$$\frac{Q_z f - Q_{z_0} f}{z - z_0} - Q_{z_0}(f\xi) = \int_{z_0}^z Q_\zeta(f\xi^2) \frac{z - \zeta}{z - z_0} d\zeta.$$

Hence, by (A3), for all $z \in \mathbb{D}_b$,

$$\begin{aligned} \left\| \frac{Q_z f - Q_{z_0} f}{z - z_0} - Q_{z_0}(f\xi) \right\|_\alpha &\leq \int_{z_0}^z \|Q_\zeta(f\xi^2)\|_\alpha \frac{|z - \zeta|}{|z - z_0|} |d\zeta| \\ &\leq M \|f\|_\alpha |z - z_0|, \end{aligned}$$

which proves that the following limit exists in $\mathcal{L}(\mathcal{B}_\alpha)$,

$$\lim_{z \rightarrow z_0} \frac{Q_z - Q_{z_0}}{z - z_0} = Q_{z_0}(\xi \cdot).$$

Notice that (A3) also implies the operator $Q_{z_0}(\xi \cdot)(f) := Q_{z_0}(\xi f)$ is in $\mathcal{L}(\mathcal{B}_\alpha)$. \square

Next proposition focus on the quasi-compactness and simplicity of $Q_z = Q_{K,z\xi}$, and is proved using arguments in [14, 1].

Proposition 2.4. *Consider a metric space \mathfrak{X} of observed Markov systems satisfying (A1)-(A4) in the range $[\alpha_1, \alpha_0] \subset (0, 1]$ with setting constants C, σ, M, b and θ .*

Given $\varepsilon > 0$ there exist $C', M' > 0$ and $0 < b_0 < b$ such that the following statement holds: for all $(K, \mu, \xi) \in \mathfrak{X}$, $z \in \mathbb{D}_{b_0}$ and $\alpha_1 \leq \alpha \leq \alpha_0$ there exist: a one dimensional subspace $E_z = E_{K,z\xi} \subset \mathcal{B}_\alpha$, a hyperplane $H_z = H_{K,z\xi} \subset \mathcal{B}_\alpha$, a number $\lambda(z) = \lambda_{K,\xi}(z) \in \mathbb{C}$, and a linear map $P_z = P_{K,z\xi} \in \mathcal{L}(\mathcal{B}_\alpha)$ such that

- (1) $\mathcal{B}_\alpha = E_z \oplus H_z$ is a Q_z -invariant decomposition,
- (2) P_z is a projection onto E_z , parallel to H_z ,
- (3) $Q_z \circ P_z = P_z \circ Q_z = \lambda(z) P_z$,
- (4) $Q_z f = \lambda(z) f$ for all $f \in E_z$,
- (5) $z \mapsto \lambda(z)$ is analytic in a neighborhood of $\overline{\mathbb{D}}_{b_0}$,
- (6) $|\lambda(z)| \geq 1 - \varepsilon$.

Furthermore, for all $f \in \mathcal{B}_\alpha$,

- (7) $\|Q_z^n f - \lambda(z)^n P_z f\|_\alpha \leq C' (\sigma + \varepsilon)^n \|f\|_\alpha$,
- (8) $\|P_z f\|_\alpha \leq C' \|f\|_\alpha$,
- (9) $\|P_z f - P_0 f\|_\alpha \leq C' |z| \|f\|_\alpha$,

and for all $z \in \mathbb{D}_{b_0}$ and $(K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2) \in \mathfrak{X}$,

- (10) $|\lambda_{K_1, \xi_1}(z) - \lambda_{K_2, \xi_2}(z)| \leq M' d((K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2))^{\frac{\theta}{2}}$.

Given $(K, \mu, \xi) \in \mathfrak{X}$, define the operators

$$P_z = P_{K, z\xi} := \frac{1}{2\pi i} \int_{\Gamma_1} R_z(w) dw \quad (3)$$

$$L_z = L_{K, z\xi} := \frac{1}{2\pi i} \int_{\Gamma_1} w R_z(w) dw \quad (4)$$

$$N_z = N_{K, z\xi} := \frac{1}{2\pi i} \int_{\Gamma_0} w R_z(w) dw \quad (5)$$

where Γ_0 and Γ_1 are the positively oriented circles

$$\Gamma_0 = \{w \in \mathbb{C} : |w| = \frac{1+2\sigma}{3}\},$$

$$\Gamma_1 = \{w \in \mathbb{C} : |w-1| = \frac{1-\sigma}{3}\},$$

and $R_z(w) = R_{K, z\xi}$ stands for the resolvent of $Q_{K, z\xi}$,

$$R_z(w) := (wI - Q_{K, z\xi})^{-1}.$$

Lemma 2.5. *Given a normed space $(\mathcal{B}, \|\cdot\|)$ and linear operators $T, T_0 \in \mathcal{L}(\mathcal{B})$,*

if T_0 is invertible with $\|T_0^{-1}\| \leq C$ and $\|T - T_0\| \leq \varepsilon$ then

- (1) *T is invertible, with $\|T^{-1}\| \leq \frac{C}{1 - C\varepsilon}$,*
- (2) *$\|T^{-1} - T_0^{-1}\| \leq \frac{C^2}{1 - C\varepsilon} \|T - T_0\|$.*

Proof. Since $T^{-1} = \sum_{n=0}^{\infty} (-1)^n (T_0^{-1} (T - T_0))^n T_0^{-1}$, we have

$$\|T^{-1}\| \leq \sum_{n=0}^{\infty} \|T_0^{-1}\|^{n+1} \|T - T_0\|^n = \frac{\|T_0^{-1}\|}{1 - \|T_0^{-1}\| \|T - T_0\|} \leq \frac{C}{1 - C\varepsilon}.$$

For (2) use the formula $T^{-1} - T_0^{-1} = -T^{-1}(T - T_0)T_0^{-1}$. \square

Lemma 2.6. *There exist constants $C_0 > 0$ and $0 < b_0 < b$, depending only on C, M, σ and b , such that for $(K, \mu, \xi) \in \mathfrak{X}$, $z \in \mathbb{D}_{b_0}$, and any of the five operators $T_z = Q_{K,z\xi}$, L_z , N_z , P_z , and $R_z(w)$ with $w \notin \text{int}(\Gamma_0) \cup \text{int}(\Gamma_1)$,*

- (1) $\|T_z\| \leq C_0$,
- (2) $\|T_z - T_0\| \leq C_0 |z|$.

Proof. First note that $\|L_0\| = \|P_0\| = 1$ and $\|N_0\| \leq C\sigma$, so that $\|Q_0\| = \|L_0 + N_0\| \leq 1 + C\sigma$. Let us go through the given operators, one at a time. Assume $0 < b_0 < b$ is small and take $z \in \mathbb{D}_{b_0}$. For $Q_{K,z\xi}$, item (1) follows from assumption (A3), taking $C_0 := M$, while (2) follows from (A3) and Proposition 2.3 with the same constant. For the operator $R_z(w)$, we have

$$\begin{aligned} R_0(w) &= w^{-1} (I - w^{-1} Q_0)^{-1} = w^{-1} \sum_{n=0}^{\infty} \frac{Q_0^n}{w^n} \\ &= w^{-1} \sum_{n=0}^{\infty} \frac{P_0}{w^n} + w^{-1} \sum_{n=0}^{\infty} \frac{N_0^n}{w^n} = \frac{P_0}{w-1} + \sum_{n=0}^{\infty} \frac{N_0^n}{w^{n+1}}. \end{aligned}$$

Notice also that $w \notin \text{int}(\Gamma_0) \cup \text{int}(\Gamma_1)$ implies $|w-1| \geq \frac{1-\sigma}{3}$ and $|w| \geq \frac{1+2\sigma}{3}$, and hence

$$\begin{aligned} \|R_0(w)\| &\leq \frac{\|P_0\|}{|w-1|} + \frac{C}{|w|} \sum_{n=0}^{\infty} \left(\frac{\sigma}{|w|} \right)^n \\ &\leq \frac{3}{1-\sigma} + \frac{3C}{1+2\sigma} \sum_{n=0}^{\infty} \left(\frac{3\sigma}{1+2\sigma} \right)^n = \frac{3+3C}{1-\sigma} =: C_1. \end{aligned}$$

Therefore, applying Lemma 2.5 to $wI - Q_z$ and $wI - Q_0$, item (1) holds with $C_2 := \frac{C_1}{1-C_1 C_0 b_0}$, while (2) holds with $C_3 := \frac{C_1^2 C_0}{1-C_1 C_0 b_0}$. Of course we have to pick $0 < b_0 < b$ small enough to make sure the denominators in constants C_2 and C_3 are both positive. For the remaining operators P_z , L_z and N_z we use the integral formulas (3), (4) and (5) to reduce to the previous case, using the same constants C_2 and C_3 as before. \square

Proof of Proposition 2.4. By Lemma 2.6 for all $|z| < b$ and $w \notin \text{int}(\Gamma_0) \cup \text{int}(\Gamma_1)$, the operator norm $\|R_z(w)\|$ is uniformly bounded. This implies that the spectrum Σ_z of $Q_{K,z\xi}$ is contained in $\text{int}(\Gamma_0) \cup \text{int}(\Gamma_1)$, and hence we can write $\Sigma_z = \Sigma_z^0 \cup \Sigma_z^1$ with $\Sigma_z^i \subset \text{int}(\Gamma_i)$, for $i = 0, 1$. By the spectral theory of bounded operators on Banach spaces, see for instance chapter IX in [19], if we denote by H_z and E_z the subspaces

of \mathcal{B}_α , respectively associated to the spectrum components Σ_z^0 and Σ_z^1 , then for all $z \in \mathbb{D}_{b_0}$, with $b_0 > 0$ small enough,

- (a) the operators Q_z , P_z , L_z and N_z commute,
- (a) $L_z f = Q_z f \in E_z$, for all $f \in E_z$,
- (b) $N_z f = Q_z f \in H_z$, for all $f \in H_z$,
- (c) $Q_z = L_z + N_z$,
- (d) $\mathcal{B}_\alpha = E_z \oplus H_z$,
- (e) P_z is the projection to E_z parallel to H_z .

For $z = 0$, the condition (A2) implies that the operator $Q_0|_{\mathcal{B}_\alpha}$ is quasi-compact and simple, with spectrum $\Sigma_0^0 \subset \mathbb{D}_\sigma$ and $\Sigma_0^1 = \{1\}$. Since 1 is a simple eigenvalue, $E_0 = \langle \mathbf{1} \rangle$ is the space of constant functions. We must have $H_0 = \{f \in \mathcal{B}_\alpha : \int f d\mu = 0\}$ because the operator Q_0 acts invariantly on this space, as a contraction with spectral radius $\leq \sigma$. Thus for all $f \in \mathcal{B}_\alpha$, $P_0 f = (\int f d\mu) \mathbf{1}$ and $N_0 f = Q_0 f - (\int f d\mu) \mathbf{1}$. Since 1 is a simple eigenvalue of Q_0 , a continuity argument implies that Σ_z^1 is a singleton, i.e., $\Sigma_z^1 = \{\lambda(z)\}$, for all $z \in \mathbb{D}_b$. It follows easily that $\dim(E_z) = 1$, and $\lambda(z) = \langle L_z \mathbf{1}, \mu \rangle / \langle P_z \mathbf{1}, \mu \rangle$. By perturbation theory, and Proposition 2.3, the function $\lambda : \mathbb{D}_{b_0} \rightarrow \mathbb{C}$ is analytic. Hence, to finish the proof of Proposition 2.4, it is now enough to establish items (6)-(10).

Take $0 < b_0 < b$ according to Lemma 2.6. Fixing a reference probability measure μ_0 on Σ , we can write, for all $z \in \mathbb{D}_b$,

$$\lambda_{K,\mu,\xi}(z) = \frac{\langle L_{K,z} \xi \mathbf{1}, \mu_0 \rangle}{\langle P_{K,z} \xi \mathbf{1}, \mu_0 \rangle}. \quad (6)$$

Notice that by Lemma 2.6, for all $(K, \mu, \xi) \in \mathfrak{X}$,

$$\langle P_{K,z} \xi \mathbf{1}, \mu_0 \rangle \geq 1 - \|P_{K,z} \xi \mathbf{1} - P_{K,0} \mathbf{1}\|_\alpha \geq 1 - C_0 b_0.$$

Hence, for all $z \in \mathbb{D}_{b_0}$,

$$\begin{aligned} |\lambda_{K,\mu,\xi}(z) - 1| &\leq \left| \frac{\langle L_{K,z} \xi \mathbf{1}, \mu_0 \rangle}{\langle P_{K,z} \xi \mathbf{1}, \mu_0 \rangle} - \frac{\langle L_{K,0} \mathbf{1}, \mu_0 \rangle}{\langle P_{K,0} \mathbf{1}, \mu_0 \rangle} \right| \\ &\leq \frac{|\langle L_{K,z} \xi \mathbf{1} - L_{K,0} \mathbf{1}, \mu_0 \rangle|}{1 - C_0 b_0} + \frac{C_0 |\langle P_{K,z} \xi \mathbf{1} - P_{K,0} \mathbf{1}, \mu_0 \rangle|}{(1 - C_0 b_0)^2} \\ &\leq \frac{C_0 b_0}{1 - C_0 b_0} + \frac{C_0^2 b_0}{(1 - C_0 b_0)^2} = O(b_0). \end{aligned}$$

Thus, given $\varepsilon > 0$ we can make $b_0 > 0$ small enough so that for all $(K, \mu, \xi) \in \mathfrak{X}$, and all $z \in \mathbb{D}_{b_0}$, $|\lambda_{K,\mu,\xi}(z) - 1| < \varepsilon$. This implies (6).

To prove (7), choose $p \in \mathbb{N}$ such that $C \sigma^p \leq (\sigma + \frac{\varepsilon}{2})^p$, and make $b_0 > 0$ small enough so that

$$p C_0^p b_0 < (\sigma + \varepsilon)^p - (\sigma + \frac{\varepsilon}{2})^p = O(\varepsilon) .$$

We have then

$$\begin{aligned} \|N_z^p\| &\leq \|N_0^p\| + \|N_z^p - N_0^p\| \\ &\leq C \sigma^p + p C_0^{p-1} \|N_z - N_0\| \leq C \sigma^p + p C_0^p b_0 \\ &\leq C \sigma^p + (\sigma + \varepsilon)^p - (\sigma + \frac{\varepsilon}{2})^p < (\sigma + \varepsilon)^p . \end{aligned}$$

It follows that for all $n \in \mathbb{N}$, $\|N_z^n\| \leq C_0^n (\sigma + \varepsilon)^n$. This proves (7) with $C' = C_0^p$.

Items (8) and (9) follow from Lemma 2.6.

To prove item (10), we claim that for all $(K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2) \in \mathfrak{X}$, $z \in \mathbb{D}_{b_0}$, $2\alpha_1 \leq \alpha \leq \alpha_0$, and $f \in \mathcal{B}_\alpha$,

$$v_{\frac{\alpha}{2}}(Q_{K_1, z\xi_1} f - Q_{K_2, z\xi_2} f) \lesssim \|f\|_\alpha \text{dist}((K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2))^{\frac{\theta}{2}} . \quad (7)$$

In fact by (B7), (B2) and (A4), we have

$$\begin{aligned} v_{\frac{\alpha}{2}}(Q_{K_1, z\xi_1} f - Q_{K_2, z\xi_2} f) &\leq v_0(Q_{K_1, z\xi_1} f - Q_{K_2, z\xi_2} f)^{\frac{1}{2}} v_\alpha(Q_{K_1, z\xi_1} f - Q_{K_2, z\xi_2} f)^{\frac{1}{2}} \\ &\lesssim \|Q_{K_1, z\xi_1} f - Q_{K_2, z\xi_2} f\|_\infty^{\frac{1}{2}} v_\alpha(Q_{K_1, z\xi_1} f - Q_{K_2, z\xi_2} f)^{\frac{1}{2}} \\ &\lesssim \|f\|_\alpha \text{dist}((K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2))^{\frac{\theta}{2}} . \end{aligned}$$

Equation (7) implies, for all $(K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2), z, \alpha$ and f as above, and all $w \notin \text{int}(\Gamma_0) \cup \text{int}(\Gamma_1)$,

$$v_{\frac{\alpha}{2}}(R_{K_1, z\xi_1}(w)f - R_{K_2, z\xi_2}(w)f) \lesssim \|f\|_\alpha \text{dist}((K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2))^{\frac{\theta}{2}} . \quad (8)$$

This follows from (7), Lemma 2.6, and the algebraic relation

$$R_{K_1, z\xi_1}(w) - R_{K_2, z\xi_2}(w) = -R_{K_1, z\xi_1}(w) \circ (Q_{K_1, z\xi_1} - Q_{K_2, z\xi_2}) \circ R_{K_2, z\xi_2}(w) .$$

Thus, integrating (3) and (4), we obtain

$$\begin{aligned} \|P_{K_1, z\xi_1} f - P_{K_2, z\xi_2} f\|_{\frac{\alpha}{2}} &\lesssim \|f\|_\alpha \text{dist}((K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2))^{\frac{\theta}{2}} , \\ \|L_{K_1, z\xi_1} f - L_{K_2, z\xi_2} f\|_{\frac{\alpha}{2}} &\lesssim \|f\|_\alpha \text{dist}((K_1, \mu_1, \xi_1), (K_2, \mu_2, \xi_2))^{\frac{\theta}{2}} . \end{aligned}$$

Finally, (10) follows from the previous inequalities and (6). \square

Remark 2.1. The condition $\alpha_1 < \frac{\alpha_0}{2}$ and the assumption (A4) are only needed to prove (10).

2.2. An abstract theorem. In this subsection we state and prove an abstract LDT theorem.

Let $(\mathcal{B}_\alpha, \|\cdot\|)_{\alpha \in [0,1]}$ be a scale of Banach algebras satisfying (B1)-(B7). Assume \mathfrak{X} is a metric space of observed Markov systems for which assumptions (A1)-(A4) hold. Take $0 < b_0 < b$ according to proposition 2.4.

Given $(K, \mu, \xi) \in \mathfrak{X}$, let $c_{K,\xi}(z) := \log \lambda_{K,\xi}(z)$, where $\lambda_{K,\xi}(z)$ denotes the maximal eigenvalue of $Q_{K,\xi}$.

Theorem 2.1. *Given $(K_0, \mu_0, \xi_0) \in \mathfrak{X}$ and $h > (c_{K_0,\xi_0})''(0)$, there exist a neighbourhood \mathcal{V} of $(K_0, \mu_0, \xi_0) \in \mathfrak{X}$, $C > 0$ and $\varepsilon_0 > 0$ such that for all $(K, \mu, \xi) \in \mathcal{V}$, $0 < \varepsilon < \varepsilon_0$, $x \in \Sigma$ and $n \in \mathbb{N}$,*

$$\mathbb{P}_x^+ \left[\left| \frac{1}{n} S_n(\xi) - \mathbb{E}_\mu(\xi) \right| \geq \varepsilon \right] \leq C e^{-\frac{\varepsilon^2}{2h} n}. \quad (9)$$

Remark 2.2. *Averaging in x , w.r.t. μ , the probabilities in theorem 2.1, we get for all $0 < \varepsilon < \varepsilon_0$, $(K, \mu, \xi) \in \mathcal{V}$ and $n \in \mathbb{N}$,*

$$\mathbb{P}_\mu^+ \left[\left| \frac{1}{n} S_n(\xi) - \mathbb{E}_\mu(\xi) \right| \geq \varepsilon \right] \leq C e^{-\frac{\varepsilon^2}{2h} n}.$$

Lemma 2.7. *For all $(K, \mu, \xi) \in \mathfrak{X}$, $n \in \mathbb{N}$, $z \in \mathbb{D}_{b_0}$ and $x \in \Sigma$,*

$$((Q_{K,z\xi})^n \mathbf{1})(x) = \mathbb{E}_x [e^{z S_n(\xi)}] = \int_{X^+} e^{z S_n(\xi)} d\mathbb{P}_x^+.$$

In particular, for all $z \in \mathbb{D}_{b_0}$,

$$\mathbb{E}_\mu((Q_{K,z\xi})^n \mathbf{1}) = \mathbb{E}_\mu [e^{z S_n(\xi)}].$$

Proof. In fact,

$$((Q_{K,z\xi})^n \mathbf{1})(x_0) = \int_{\Sigma^n} e^{z \sum_{j=1}^n \xi(x_j)} \prod_{j=0}^{n-1} K(x_j, dx_{j+1}) = \mathbb{E}_{x_0} [e^{z S_n(\xi)}].$$

Averaging this relation in x_0 w.r.t. μ we derive the second identity. \square \square

Next proposition shows that $c_{K,\xi}(z)$ is a limit cumulant generating function of the process $\{S_n(\xi)\}_{n \geq 0}$. Moreover it says that the parameters C and δ_n in definition 1.16 can be chosen uniformly in \mathfrak{X} .

Proposition 2.8. *There exist $C_1 > 0$ and a sequence δ_n converging geometrically to 0 such that for all $(K, \mu, \xi) \in \mathfrak{X}$, $z \in \mathbb{D}_{b_0}(0)$, $x \in \Sigma$ and $n \in \mathbb{N}$*

$$|n \log \lambda_{K,\xi}(z) - \log \mathbb{E}_x [e^{z S_n(\xi)}]| \leq C_1 |z| + \delta_n.$$

Proof. We will use the notation of Proposition 2.4, choosing $\varepsilon > 0$ small enough so that $\sigma + \varepsilon < 1 - \varepsilon$. By Lemma 2.7, $(Q_z^n \mathbf{1})(x) = \mathbb{E}_x [e^{z S_n(\xi)}]$. By Lemma 2.6 there exists $B > 0$ such that for all $z \in \mathbb{D}_{b_0}(0)$, $\|P_z - I\|_\alpha \leq B|z|$. Hence

$$\begin{aligned} |\mathbb{E}_x [e^{z S_n(\xi)}] - \lambda_{K,\xi}(z)^n| &\leq |(Q_z^n \mathbf{1})(x) - \lambda_{K,\xi}(z)^n| \\ &\leq \|Q_z^n \mathbf{1} - \lambda_{K,\xi}(z)^n P_z \mathbf{1}\|_\alpha + \lambda_{K,\xi}(z)^n \|\mathbf{1} - P_z \mathbf{1}\|_\alpha \\ &= \|N_z^n \mathbf{1}\|_\alpha + \lambda_{K,\xi}(z)^n \|\mathbf{1} - P_z \mathbf{1}\|_\alpha \\ &\leq C(\sigma + \varepsilon)^n + B|z| \lambda_{K,\xi}(z)^n. \end{aligned}$$

Thus

$$\begin{aligned} |\log \mathbb{E}_x [e^{t S_n(\xi)}] - n \log \lambda_{K,\xi}(z)| &= |\log \mathbb{E}_x [e^{t S_n(\xi)}] - \log \lambda_{K,\xi}(z)^n| \\ &\leq \frac{|\mathbb{E}_x [e^{t S_n(\xi)}] - \lambda_{K,\xi}(z)^n|}{\min\{\lambda_{K,\xi}(z)^n, \mathbb{E}_x [e^{t S_n(\xi)}]\}} \\ &\leq \frac{B|z| \lambda_{K,\xi}(z)^n + C(\sigma + \varepsilon)^n}{(1 - B|z|) \lambda_{K,\xi}(z)^n - C(\sigma + \varepsilon)^n} \\ &\leq \frac{B|z| + \delta_n}{1 - B|z| - \delta_n} \leq 2(B|z| + \delta_n), \end{aligned}$$

where $\delta_n := C \frac{(\sigma + \varepsilon)^n}{\lambda_{K,\xi}(z)^n} \leq C \left(\frac{\sigma + \varepsilon}{1 - \varepsilon} \right)^n$ converges geometrically to zero. \square

\square

theorem 2.1. Combine Proposition 2.8 with Corollary 1.8. \square \square

3. THE PROOF OF LDT ESTIMATES

We prove here the base-LDT and uniform fiber-LDT estimates for irreducible cocycles over mixing Markov shifts. These results follow from the abstract Theorem 2.1.

3.1. Base LDT estimates. To deduce theorem 1.2 from theorem 2.1 we specify the data $(\mathcal{B}_\alpha, \|\cdot\|_\alpha)$ and \mathfrak{X} , and check the validity of the assumptions (B1)-(B7) and (A1)-(A4).

Consider a strongly mixing Markov system (K, μ) on the compact metric space Σ . Let $X^- = \Sigma^{\mathbb{Z}_0^-}$ be the space of sequences in Σ indexed in the set \mathbb{Z}_0^- of non-positive integers. Since \mathbb{Z}_0^- is countable, the product X^- is a compact metrizable topological space. We denote by \mathcal{F} its Borel σ -field. The kernel K on Σ induces another Markov kernel \tilde{K} on X^- defined by

$$\tilde{K}_{(\dots, x_{-1}, x_0)} := \int_{\Sigma} \delta_{(\dots, x_{-1}, x_0, x_1)} K(x_0, dx_1).$$

Let \mathbb{P}_μ^- denote the Kolmogorov extension of (K, μ) , which is also the unique \tilde{K} -stationary measure. Theorem 2.1 will be applied to the Markov system $(\tilde{K}, \mathbb{P}_\mu^-)$.

Consider the spaces $\mathcal{H}_\alpha(X^-)$ introduced in definition (1.11). Its functions can be regarded as measurable functions on X^- . They form the scale of Banach algebras satisfying (B1)-(B7). See proposition 2.2. The metric space (X^-, \tilde{d}) has diameter 1 but is not compact, as noticed after the definition (1) of the distance \tilde{d} . Hence, formally, the claim above is not a direct consequence of proposition 2.2. Properties (B1), (B3) and (B4) follow from proposition 1.4. For $\alpha = 0$, the seminorm v_0 measures the variation of f . Hence $\mathcal{H}_0(X) = L^\infty(X)$, while the norm $\|\cdot\|_0$ is equivalent to $\|\cdot\|_\infty$. This proves (B2). The remaining properties, (B5)-(B7), can be proved as in proposition 2.2.

Fix $0 < \alpha_0 \leq 1$ and $0 < L < +\infty$ and consider the space \mathfrak{X} of observed Markov systems $(\tilde{K}, \mathbb{P}_\mu^-, \xi)$ over the fixed Markov system $(\tilde{K}, \mathbb{P}_\mu^-)$, with $\xi \in \mathcal{H}_{\alpha_0}(X^-)$ and $\|\xi\|_{\alpha_0} \leq L$. This space is identified with a subspace of $\mathcal{H}_{\alpha_0}(X^-)$, and endowed with the corresponding norm distance.

The kernel \tilde{K} determines the Markov operator $Q_{\tilde{K}} : L^\infty(X^-) \rightarrow L^\infty(X^-)$,

$$(Q_{\tilde{K}}f)(\dots, x_{-1}, x_0) := \int_{\Sigma} f(\dots, x_{-1}, x_0, x_1) K(x_0, dx_1) .$$

This operator acts continuously on $\mathcal{H}_\alpha(X^-)$.

Proposition 3.1. *For all $f \in \mathcal{H}_\alpha(X^-)$ and $n \in \mathbb{N}$,*

- (1) $\|(Q_{\tilde{K}})^n f\|_\infty \leq \|f\|_\infty$,
- (2) $v_\alpha((Q_{\tilde{K}})^n f) \leq \max\{2\|(Q_{\tilde{K}})^n f\|_\infty, 2^{-n\alpha}v_\alpha(f)\}$.

Proof. We shall write $Q = Q_{\tilde{K}}$. Since $\int_{\Sigma} K(x_0, dx_1) = 1$, the first inequality follows. For the second, notice that if $k \geq 1$ then $v_k(Q^n f) \leq v_{k+n}(f)$. Indeed, for $x = (x_n)_{n \leq 0}$ and $x' = (x'_n)_{n \leq 0}$ in X^- such that $\tilde{d}(x, x') \leq 2^{-k}$ with $k \geq 1$, we have $x_0 = x'_0$. Thus

$$\begin{aligned} & |(Q^n f)(\dots, x_{-1}, x_0) - (Q^n f)(\dots, x'_{-1}, x'_0)| \\ & \leq \int_{\Sigma^n} |f(\dots, x_0, x_1, \dots, x_n) - f(\dots, x'_0, x_1, \dots, x_n)| \prod_{j=0}^{n-1} K(x_j, dx_{j+1}) \\ & \leq v_{k+n}(f) \int_{\Sigma^n} \prod_{j=0}^{n-1} K(x_j, dx_{j+1}) = v_{k+n}(f) , \end{aligned}$$

and taking the sup in $x, x' \in X^-$ such that $\tilde{d}(x, x') \leq 2^{-k}$, the inequality $v_k(Q^n f) \leq v_{k+n}(f)$ follows. Hence, for $k \geq 1$,

$$2^{\alpha k} v_k(Q^n f) = 2^{-n\alpha} (2^{\alpha(k+n)} v_{k+n}(f)) \leq 2^{-n\alpha} v_\alpha(f) .$$

For $k = 0$ notice that $v_0(Q^n f)$ is the variation of $Q^n f$. Thus $v_0(Q^n f) \leq 2 \|Q^n f\|_\infty$. Taking the sup in $k \in \mathbb{N}$, item (2) follows. \square

Next proposition shows that \mathfrak{X} satisfies (A2) with range $[\alpha_1, \alpha]$ for any given $0 < \alpha_1 \leq \alpha$. The setting constants $C > 0$ and $0 < \sigma < 1$ depend on the number α_1 .

Proposition 3.2. *If (K, μ) is strongly mixing, then given $0 < \alpha_1 < \alpha_0$ there are constants $C > 0$ and $0 < \sigma < 1$ such that for all $\alpha_1 \leq \alpha \leq \alpha_0$, $Q_{\tilde{K}} : \mathcal{H}_\alpha(X^-) \rightarrow \mathcal{H}_\alpha(X^-)$ is quasi-compact and simple with spectral constants C and σ , i.e., for all $f \in \mathcal{H}_\alpha(X^-)$,*

$$\|(Q_{\tilde{K}})^n f - \langle f, \mathbb{P}_\mu^- \rangle \mathbf{1}\|_\alpha \leq C \sigma^n \|f\|_\alpha .$$

Proof. Given a function $f \in \mathcal{H}_\alpha(X^-)$, denote by $f_k : X^- \rightarrow \mathbb{C}$ the following function

$$f_k(\dots, x_0) := \int_{X^-} f(\dots, x_{-k}, \dots, x_0) d\mathbb{P}_\mu^-(\dots, x_{-k}) .$$

Note that if \mathcal{F}_k^- is the sub σ -field of \mathcal{F}^- generated by the cylinders in the coordinates $x_{-k+1}, \dots, x_{-1}, x_0$, we have $f_k = \mathbb{E}_\mu^-(f | \mathcal{F}_k^-)$, and in particular $\mathbb{E}_\mu^-(f_k) = \mathbb{E}_\mu^-(f)$, for all $k \in \mathbb{N}$. By definition of f_k ,

$$\|Q^n(f - f_k)\|_\infty \leq \|f - f_k\|_\infty \leq v_k(f) \leq 2^{-\alpha k} v_\alpha(f) . \quad (10)$$

Because (K, μ) is strongly mixing, there are constants $C > 0$ and $0 < \rho < 1$ such that for any function $h \in L^\infty(\Sigma)$ with $\int_\Sigma h d\mu = 0$,

$$\left| \int_\Sigma h(y) K^n(x, dy) \right| \leq C \rho^n \|h\|_\infty .$$

Now, if $h \in L^\infty(X^-)$ is a function with zero average, i.e., $\mathbb{E}_\mu^-(h) = 0$, which depends only on the first coordinate x_0 , then $Q^n h$ also depends only on the first coordinate, and is given by

$$(Q^n h)(\dots, x_0) := \int_\Sigma h(y) K^n(x_0, dy) .$$

Hence

$$\|Q^n h\|_\infty \leq C \rho^n \|h\|_\infty . \quad (11)$$

We claim that $h = Q^k(f_k - \mathbb{E}_\mu^-(f) \mathbf{1})$ is a function with zero average that depends only on the first coordinate. The first part of claim follows because Q preserves averages and, as remarked above, $\mathbb{E}_\mu^-(f_k) = \mathbb{E}_\mu^-(f)$. For the second part notice two things: first Q ‘preserves’ functions that

depend only on the first coordinate x_0 ; second, Q maps a function f that depends only on the coordinates $x_{-k}, \dots, x_{-1}, x_0$ to a function that depends only on the coordinates $x_{-k+1}, \dots, x_{-1}, x_0$, in other words Qf loses dependence in x_{-k} . Therefore, from (11)

$$\begin{aligned} \|Q^n(f_k - \mathbb{E}_\mu^-(f) \mathbf{1})\|_\infty &= \|Q^{n-k}h\|_\infty \leq C \rho^{n-k} \|h\|_\infty \\ &\leq C \rho^{n-k} \|Q^k(f_k - \mathbb{E}_\mu^-(f) \mathbf{1})\|_\infty \\ &\leq C \rho^{n-k} \|f_k - \mathbb{E}_\mu^-(f) \mathbf{1}\|_\infty \leq 2C \rho^{n-k} \|f\|_\infty \end{aligned} \quad (12)$$

Setting $\sigma = \max\{2^{-\frac{\alpha-1}{2}}, \sqrt{\rho}\}$ we have $0 < \sigma < 1$. From the inequalities (10) and (12), with $k = n/2$, we have

$$\begin{aligned} \|Q^n f - \mathbb{E}_\mu^-(f) \mathbf{1}\|_\infty &\leq \|Q^n(f - f_k)\|_\infty + \|Q^n(f_k - \mathbb{E}_\mu^-(f) \mathbf{1})\|_\infty \\ &\leq 2^{-\alpha \frac{n}{2}} v_\alpha(f) + 2C \rho^{\frac{n}{2}} \|f\|_\infty \\ &\leq \sigma^n v_\alpha(f) + 2C \sigma^n \|f\|_\infty. \end{aligned}$$

On the other hand, by item (2) of Proposition 3.1,

$$\begin{aligned} v_\alpha(Q^n f - \mathbb{E}_\mu^-(f) \mathbf{1}) &= v_\alpha(Q^n(f - \mathbb{E}_\mu^-(f) \mathbf{1})) \\ &\leq \max\{\|Q^n f - \mathbb{E}_\mu^-(f) \mathbf{1}\|_\infty, 2^{-n\alpha} v_\alpha(f)\} \\ &\leq \max\{\sigma^n v_\alpha(f) + 2C \sigma^n \|f\|_\infty, \sigma^{2n} v_\alpha(f)\} \\ &\leq \sigma^n v_\alpha(f) + 2C \sigma^n \|f\|_\infty. \end{aligned}$$

Thus, for all $f \in \mathcal{H}_\alpha(X^-)$,

$$\|Q^n f - \mathbb{E}_\mu^-(f) \mathbf{1}\|_\alpha \leq 4C \sigma^n \|f\|_\alpha,$$

which proves the proposition. \square

3.2. Fiber LDT estimates. In this subsection we use theorem 2.1 to establish the fiber LDT theorem 1.3. First we specify the data $(\mathcal{B}_\alpha, \|\cdot\|_\alpha)$ and the metric space \mathfrak{X} . Then we check that assumptions (B1)-(B7) and (A1)-(A4) hold up.

Consider the space $\mathcal{B}_m^\infty(K)$ of random cocycles over a Markov system (K, μ) . For each cocycle $A \in \mathcal{B}_m^\infty(K)$ we define a Markov kernel on $\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m)$ by

$$K_A(x, y, p) := \int_{\Sigma} \delta_{(y, z, A(y, z)p)} K(y, dz). \quad (13)$$

We will see that (c.f. corollary 3.13), under the assumptions of theorem 1.3, this kernel admits a unique K_A -stationary probability measure μ_A in $\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m)$.

For each $A \in \mathcal{B}_m^\infty(K)$ consider the observable $\xi_A : \Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m) \rightarrow \mathbb{R}$

$$\xi_A(x, y, p) := \log \|A(x, y) p\|. \quad (14)$$

We can now introduce the metric space of observed Markov systems

$$\mathfrak{X} := \{ (K_A, \mu_A, \pm \xi_A) : A \in \mathcal{B}_m^\infty(K), A \text{ irreducible}, L_1(A) > L_2(A) \}.$$

This space is identified with a subspace of $\mathcal{B}_m^\infty(K)$, and endowed with the distance

$$\text{dist}((K_A, \mu_A, \xi_A), (K_B, \mu_B, \xi_B)) := d_\infty(A, B).$$

Next we define the scale of Banach algebras. Consider the following projective distance (see [5, formula (1.3)])

$$\delta(\hat{p}, \hat{q}) := \frac{\|p \wedge q\|}{\|p\| \|q\|},$$

where $p \in \hat{p}$ and $q \in \hat{q}$. Given $0 \leq \alpha \leq 1$ and $f \in L^\infty(\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m))$, let

$$\|f\|_\alpha := v_\alpha(f) + \|f\|_\infty, \quad (15)$$

$$v_\alpha(f) := \sup_{\substack{x, y \in \Sigma \\ p \neq q}} \frac{|f(x, y, p) - f(x, y, q)|}{\delta(p, q)^\alpha}. \quad (16)$$

Definition 3.1. Consider the normed space $\mathcal{H}_\alpha(\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m))$ of all functions $f \in L^\infty(\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m))$ such that $v_\alpha(f) < +\infty$, endowed with the norm (15).

Proposition 3.3. The family of spaces $\mathcal{H}_\alpha(\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m))$ is a scale of Banach algebras satisfying (B1)-(B7).

Proof. (B1) holds by definition of the Hölder norm $\|\cdot\|_\alpha$. For (B2) notice that $v_0(f)$ measures the maximum oscillation of f on the projective fibers, and hence $v_0(f) \leq 2\|f\|_\infty$. Property (B3) is obvious. Assumption (B4) is a consequence of the inequality

$$v_\alpha(fg) \leq \|f\|_\infty v_\alpha(g) + \|g\|_\infty v_\alpha(f), \quad f, g \in L^\infty(\Sigma).$$

The monotonicity properties (B5) and (B6) are straightforward to check. The assumption (B7) follows from the convexity of the function $\alpha \mapsto \log v_\alpha(f)$, whose proof is analogous to that of proposition 2.2. \square

Definition 3.2. We define $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ to be the subspace of functions $f(x, y, p)$ in $\mathcal{H}_\alpha(\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m))$ that do not depend on the first coordinate x .

This subspace is clearly a closed sub-algebra of $\mathcal{H}_\alpha(\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m))$. Therefore,

Proposition 3.4. *The family $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ is a scale of Banach sub-algebras satisfying (B1)-(B7).*

Given $A \in \mathcal{B}_m^\infty(K)$, consider the linear transformation $Q_A : L^\infty(\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m)) \rightarrow L^\infty(\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m))$ defined by

$$(Q_A f)(x, y, p) := \int_{\Sigma} f(y, z, A(y, z)p) K(y, dz) . \quad (17)$$

This is the Markov operator associated with the kernel (13).

Assumption (A1) follows from the definition of \mathfrak{X} .

Since $(Q_A f)(x, y, p)$ does not depend on the coordinate x , the Markov operator Q_A leaves invariant the subspace of functions $f(x, y, p)$ that are constant in x . Next, we are going to see that Q_A acts invariantly on the subspace $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$.

Given $A \in \mathcal{B}_m^\infty(K)$ and $0 < \alpha \leq 1$, define for all $n \in \mathbb{N}$,

$$\kappa_\alpha^n(A) := \sup_{x \in \Sigma, p \neq q} \mathbb{E}_x \left[\left(\frac{\delta(A^{(n)}p, A^{(n)}q)}{\delta(p, q)} \right)^\alpha \right] \in [0, +\infty] \quad (18)$$

Lemma 3.5. *Let $A \in \mathcal{B}_m^\infty(K)$ and $n \in \mathbb{N}$.*

- (a) $\|A^{(\pm n)}\|_\infty \leq \max\{\|A\|_\infty, \|A^{-1}\|_\infty\}^n$.
- (b) $\|A^{(n)} - B^{(n)}\|_\infty \leq n \max\{\|A\|_\infty, \|B\|_\infty\}^{n-1} \|A - B\|_\infty$.

Proof. Item (a) is straightforward. To prove (b), we use the formula

$$A^{(n)} - B^{(n)} = \sum_{j=0}^{n-1} (A^{(j)} \circ T^{n-j})(A \circ T^{n-1-j} - B \circ T^{n-1-j}) B^{(n-1-j)} .$$

□

The following lemma highlights the importance of this quantity.

Lemma 3.6. *Given $A \in \mathcal{B}_m^\infty(K)$, $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ and $n \in \mathbb{N}$,*

$$v_\alpha(Q_A^n f) \leq \kappa_\alpha^n(A) v_\alpha(f) .$$

Proof. For any $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$, and $(x_0, p) \in \Sigma \times \mathbb{P}(\mathbb{R}^m)$,

$$\begin{aligned} (Q_A^n f)(x_0, p) &= \int_{\Sigma^n} f(x_n, A(x_{n-1}, x_n) \dots A(x_0, x_1)p) \prod_{j=0}^{n-1} K(x_j, dx_{j+1}) \\ &= \mathbb{E}_{x_0} [f(e_n, A^{(n)}p)] . \end{aligned}$$

Hence

$$\begin{aligned}
v_\alpha(Q_A^n f) &= \sup_{x \in \Sigma, p \neq q} \frac{|\mathbb{E}_x [f(e_n, A^{(n)}p) - f(e_n, A^{(n)}q)]|}{\delta(p, q)^\alpha} \\
&\leq \sup_{x \in \Sigma, p \neq q} \frac{\mathbb{E}_x [|f(e_n, A^{(n)}p) - f(e_n, A^{(n)}q)|]}{\delta(p, q)^\alpha} \\
&\leq v_\alpha(f) \sup_{x \in \Sigma, p \neq q} \mathbb{E}_x \left[\left(\frac{\delta(A^{(n)}p, A^{(n)}q)}{\delta(p, q)} \right)^\alpha \right] \\
&= v_\alpha(f) \kappa_\alpha^n(A) .
\end{aligned}$$

□

Lemma 3.7. *The sequence $\{\kappa_\alpha^n(A)\}_{n \geq 0}$ is sub-multiplicative, i.e.,*

$$\kappa_\alpha^{n+\ell}(A) \leq \kappa_\alpha^n(A) \kappa_\alpha^\ell(A) \quad \text{for } n, \ell \in \mathbb{N} .$$

In particular,

$$\lim_{n \rightarrow +\infty} \kappa_\alpha^n(A)^{1/n} = \inf \{ \kappa_\alpha^n(A)^{1/n} : n \in \mathbb{N} \} .$$

Proof. Let us write $M_n = A^{(n)}$. Given $x \in \Sigma$ and $p \neq q$ in $\mathbb{P}(\mathbb{R}^m)$,

$$\begin{aligned}
&\mathbb{E}_x \left[\left(\frac{\delta(M_{n+m}p, M_{n+m}q)}{\delta(p, q)} \right)^\alpha \right] \leq \\
&\leq \mathbb{E}_x \left[\left(\frac{\delta((M_n \circ T^m)M_mp, (M_n \circ T^m)M_mq)}{\delta(M_mp, M_mq)} \right)^\alpha \left(\frac{\delta(M_mp, M_mq)}{\delta(p, q)} \right)^\alpha \right] \\
&\leq \kappa_\alpha^m \mathbb{E}_x \left[\left(\frac{\delta((M_n \circ T^m)M_mp, (M_n \circ T^m)M_mq)}{\delta(M_mp, M_mq)} \right)^\alpha \right] \\
&\leq \kappa_\alpha^m \sup_{p \neq q} \mathbb{E}_{K^m(x, \cdot)} \left[\left(\frac{\delta(M_n p, M_n q)}{\delta(p, q)} \right)^\alpha \right] \leq \kappa_\alpha^m \kappa_\alpha^n ,
\end{aligned}$$

and taking the sup we get $\kappa_\alpha^{n+m} \leq \kappa_\alpha^n \kappa_\alpha^m$.

□

These constants become finite provided α is small enough.

Lemma 3.8. *Given $A \in \mathcal{B}_m^\infty(K)$ and $n \in \mathbb{N}$ for all $0 < \alpha \leq \frac{1}{4n}$,*

$$\kappa_\alpha^n(A) \leq \max\{\|A\|_\infty, \|A^{-1}\|_\infty\} .$$

Proof. We write as before $M_n = A^{(n)}$. Recall that given $M \in \text{GL}(m, \mathbb{R})$, the quantity $\ell(M) := \max\{\log\|M\|, \log\|M^{-1}\|\}$ is sub-multiplicative, in the sense that for any matrices $M_1, M_2 \in \text{GL}(m, \mathbb{R})$, $\ell(M_1 M_2) \leq \ell(M_1) + \ell(M_2)$. By [5, Lemma 3.26], given $x \in \Sigma$, and $p \neq q$ in $\mathbb{P}(\mathbb{R}^m)$,

$$\mathbb{E}_x \left[\left(\frac{\delta(M_n p, M_n q)}{\delta(p, q)} \right)^\alpha \right] = \mathbb{E}_x \left[\exp \left(\alpha \log \frac{\delta(M_n p, M_n q)}{\delta(p, q)} \right) \right] \leq \mathbb{E}_x [e^{4\alpha \ell(M_n)}] .$$

If $0 < \alpha \leq \frac{1}{4n}$, setting $c := \max\{\log\|A\|_\infty, \log\|A^{-1}\|_\infty\}$

$$\mathbb{E}_x \left[e^{4\alpha \ell(M_n)} \right] \leq e^{4n\alpha c} \leq e^c = \max\{\|A\|_\infty, \|A^{-1}\|_\infty\}.$$

Hence, taking the sup in x and $p \neq q$ we obtain $\kappa_\alpha^n \leq \max\{\|A\|_\infty, \|A^{-1}\|_\infty\}$. \square

By the previous lemmas the operator Q_A leaves the subspace $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ invariant, for all small enough $\alpha > 0$. To prove that Q_A is quasi-compact and simple all hypothesis of theorem 1.3 are essential. The irreducibility and gap assumptions are used in the following lemmas.

Lemma 3.9. *Given $A \in \mathcal{B}_m^\infty(K)$ such that $(K_A, \mu_A, \xi_A) \in \mathfrak{X}$*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \mathbb{E}_x(\log\|A^{(n)}p\|) = L_1(A),$$

with uniform convergence in $(x, p) \in \Sigma \times \mathbb{P}(\mathbb{R}^m)$.

Proof. See Lemma 3.1 in [1]. \square

Lemma 3.10. *Given $A \in \mathcal{B}_m^\infty(K)$ such that $(K_A, \mu_A, \xi_A) \in \mathfrak{X}$, there exists $n \in \mathbb{N}$ such that for all $x \in \Sigma$ and $p \neq q$ in $\mathbb{P}(\mathbb{R}^m)$,*

$$\mathbb{E}_x \left[\log \frac{\delta(A^{(n)}p, A^{(n)}q)}{\delta(p, q)} \right] \leq -1.$$

Proof. We write $M_n = A^{(n)}$. Given $x \in \Sigma$ and $p \neq q$ in $\mathbb{P}(\mathbb{R}^m)$,

$$\begin{aligned} \frac{1}{n} \mathbb{E}_x \left[\log \frac{\delta(M_n p, M_n q)}{\delta(p, q)} \right] &\leq \frac{1}{n} \mathbb{E}_x \left[\log \frac{\|(M_n p) \wedge (M_n q)\|}{\|M_n p\| \|M_n q\|} \frac{\|p\| \|q\|}{\|p \wedge q\|} \right] \\ &\leq \frac{1}{n} \mathbb{E}_x \left[\log \frac{\|(M_n p) \wedge (M_n q)\|}{\|p \wedge q\|} \frac{\|p\|}{\|M_n p\|} \frac{\|q\|}{\|M_n q\|} \right] \\ &\leq \frac{1}{n} \mathbb{E}_x [\log\|\wedge_2 A^{(n)}\|] - \frac{1}{n} \mathbb{E}_x [\log\|A^{(n)}p\|] - \frac{1}{n} \mathbb{E}_x [\log\|A^{(n)}q\|], \end{aligned}$$

and the right hand side converges to $L_1 + L_2 - 2L_1 = L_2 - L_1 < 0$. By Lemma 3.9, we have

$$\limsup_{n \rightarrow +\infty} \sup_{x \in \Sigma, p \neq q} \frac{1}{n} \mathbb{E}_x \left[\log \frac{\delta(M_n p, M_n q)}{\delta(p, q)} \right] \leq L_2 - L_1 < 0.$$

Hence taking n large enough such that $n(L_2 - L_1) < -1$ the Lemma follows. \square

Proposition 3.11. *Given $A \in \mathcal{B}_m^\infty(K)$ such that $(K_A, \mu_A, \xi_A) \in \mathfrak{X}$, there exists a neighborhood \mathcal{V} of A in $\mathcal{B}_m^\infty(K)$, and there are constants $0 < \alpha_1 < \frac{\alpha_0}{2} < \alpha_0$, $C > 0$ and $0 < \sigma < 1$ such that*

$$v_\alpha(Q_B^n f) \leq C \sigma^n v_\alpha(f),$$

for all $B \in \mathcal{V}$, $\alpha \in [\alpha_1, \alpha_0]$, $n \in \mathbb{N}$ and $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$.

Proof. We begin deriving a modulus of continuity for $B \mapsto \kappa_\alpha^n(B)$.

Fix a neighborhood \mathcal{V} of A in $\mathcal{B}_m^\infty(K)$ such that for all $B \in \mathcal{V}$, $\|B\|_\infty \leq C$ and $\|B^{-1}\|_\infty \leq C$. By Lemma 3.5 $\|B^{(\pm n)}\|_\infty \leq C^n$ for all $B \in \mathcal{V}$ and $n \in \mathbb{N}$. Thus, by [5, Lemma 3.27] and Lemma 3.5 (b), there exists a polynomial expression $C(g_1, g_2)$, with degree < 11 in the variables $\|g_1\|$, $\|g_2\|$, $\|g_1^{-1}\|$ and $\|g_2^{-1}\|$, such that

$$\begin{aligned} |\kappa_\alpha^n(A) - \kappa_\alpha^n(B)| &\leq \sup_{x \in \Sigma, p \neq q} \mathbb{E}_x \left[\left| \left(\frac{\delta(A^{(n)}p, A^{(n)}q)}{\delta(p, q)} \right)^\alpha - \left(\frac{\delta(B^{(n)}p, B^{(n)}q)}{\delta(p, q)} \right)^\alpha \right| \right] \\ &\leq \alpha C(A^{(n)}, B^{(n)}) \|A^{(n)} - B^{(n)}\|_\infty \leq \alpha C^{11n} \|A^{(n)} - B^{(n)}\|_\infty \leq \alpha n C^{12n-1} \|A - B\|_\infty. \end{aligned}$$

Let $M_n = A^{(n)}$. We claim that for some $n_0 \in \mathbb{N}$ and $0 < \alpha_0 \leq 1$ small enough, $\kappa_{\alpha_0}^{n_0}(A) < 1$. We will make use the following inequality

$$e^x \leq 1 + x + \frac{x^2}{2} e^{|x|}.$$

Choose $n_0 \in \mathbb{N}$ as given by Lemma 3.10. For all $x \in \Sigma$, $p \neq q$ in $\mathbb{P}(\mathbb{R}^m)$,

$$\begin{aligned} \mathbb{E}_x \left[\left(\frac{\delta(M_{n_0}p, M_{n_0}q)}{\delta(p, q)} \right)^\alpha \right] &= \mathbb{E}_x \left[\exp \left(\alpha \log \frac{\delta(M_{n_0}p, M_{n_0}q)}{\delta(p, q)} \right) \right] \\ &\leq \mathbb{E}_x \left[1 + \alpha \log \frac{\delta(M_{n_0}p, M_{n_0}q)}{\delta(p, q)} + \frac{\alpha^2}{2} \log^2 \frac{\delta(M_{n_0}p, M_{n_0}q)}{\delta(p, q)} \left(\frac{\delta(M_{n_0}p, M_{n_0}q)}{\delta(p, q)} \right)^\alpha \right] \\ &\leq 1 - \alpha + \frac{\alpha^2}{2} \mathbb{E}_x [16 \ell(M_{n_0})^2 \exp(\alpha \ell(M_{n_0}))] \leq 1 - \alpha + O(\alpha^2). \end{aligned}$$

The last inequality follows because $\mathbb{E}_x [16 \ell(M_{n_0})^2 \exp(\alpha \ell(M_{n_0}))]$ is finite and uniformly bounded in x and $0 < \alpha \leq 1$ by the constant $16 n_0^2 (\log C)^2 C^{n_0 \alpha}$.

Taking $\alpha > 0$ sufficiently small the right-hand-side above is less than 1, which shows that $\kappa_\alpha^{n_0}(A) < 1$. Hence, we can choose $0 < \alpha_1 < \frac{\alpha_0}{2}$ and $0 < \rho < 1$ such that for all $\alpha_1 \leq \alpha \leq \alpha_0$, $\kappa_\alpha^{n_0}(A) \leq \rho$.

Next, we extend this inequality to all cocycles $B \in \mathcal{V}$.

Pick $\rho' \in]\rho, 1[$ and choose $\delta > 0$ such that $\alpha_0 n_0 C^{12n_0-1} \delta < \rho' - \rho$. Make the neighborhood \mathcal{V} small enough so that $\|A - B\|_\infty < \delta$ for all $B \in \mathcal{V}$. Then, using the modulus of continuity for $\kappa_\alpha^n(B)$, for all $B \in \mathcal{V}$ and $\alpha_1 \leq \alpha \leq \alpha_0$,

$$|\kappa_\alpha^{n_0}(A) - \kappa_\alpha^{n_0}(B)| < \rho' - \rho,$$

which implies

$$\kappa_\alpha^{n_0}(B) \leq \kappa_\alpha^{n_0}(A) + |\kappa_\alpha^{n_0}(A) - \kappa_\alpha^{n_0}(B)| < \rho'.$$

By Lemma 3.8, $\kappa_\alpha^j(B) \leq C$ for all $B \in \mathcal{V}$, $0 < \alpha \leq \frac{1}{4n_0}$ and $0 \leq j \leq n_0$. Shrinking if necessary the constants α_1 and α_0 above, we

may assume that $\alpha_0 \leq \frac{1}{4n_0}$. Thus, because the sequence $\{\kappa_\alpha^n(B)\}_{n \geq 0}$ is sub-multiplicative, letting $\sigma = (\rho')^{1/n_0}$ we have $\kappa_\alpha^n(B) \leq C \sigma^n$ for all $B \in \mathcal{V}$, $n \in \mathbb{N}$ and $\alpha_1 \leq \alpha \leq \alpha_0$. The proposition follows then from the inequality proven in Lemma 3.6. \square

Next proposition implies (A2).

Proposition 3.12. *Given $A \in \mathcal{B}_m^\infty(K)$ such that $(K_A, \mu_A, \xi_A) \in \mathfrak{X}$, there exist a neighborhood \mathcal{V} of A in $\mathcal{B}_m^\infty(K)$, a range $0 < \alpha_1 < \frac{\alpha_0}{2} < \alpha_0 \leq 1$ and there are constants $C > 0$ and $0 < \sigma < 1$ such that for all $B \in \mathcal{V}$, $\alpha \in [\alpha_1, \alpha_0]$ and $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$,*

$$\|Q_B^n f - \langle f, \mu_B \rangle \mathbf{1}\|_\alpha \leq C \sigma^n \|f\|_\alpha.$$

Proof. The argument below is an adaptation of the proof of Theorem 3.7 in [1].

Take the neighbourhood \mathcal{V} , and the constants $\alpha_0 > 0$, $C > 0$ and $0 < \sigma < 1$ given by Proposition 3.11. Enlarging the constants $C > 0$ and $0 < \sigma < 1$ we can assume that the conditions of definition 1.6 are also satisfied with $\rho = \sigma$. By Lemma 3.6, given $B \in \mathcal{V}$ and any K_B -stationary measure ν_B ,

$$v_\alpha(Q_B^n f - \langle f, \nu_B \rangle \mathbf{1}) = v_\alpha(Q_B^n f) = v_\alpha(f) \kappa_\alpha^n(B) \leq C \sigma^n \|f\|_\alpha.$$

Hence it is now enough to prove that

$$\|Q_B^n f - \langle f, \nu_B \rangle \mathbf{1}\|_\infty \leq C \sigma^n \|f\|_\alpha.$$

We define four families of transformations

$$T_{B,n,m}^{(i)} : L^\infty(\Sigma \times \mathbb{P}(\mathbb{R}^m)) \rightarrow L^\infty(\Sigma \times \mathbb{P}(\mathbb{R}^m)) \quad i = 0, 1, 2, 3,$$

depending on $B \in \mathcal{V}$, and $n \geq m$, $n, m \in \mathbb{N}$, which act continuously on the scale of Banach spaces $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ with $0 < \alpha \leq \alpha_0$.

- $(T_{B,n}^{(0)} f)(x, p) := (Q_B^n f)(x, p) = \mathbb{E}_x [f(e_n, B^{(n)} p)]$.
- $(T_{B,n,m}^{(1)} f)(x, p) := \mathbb{E}_x [f(e_n, (B^{(m)} \circ T^{n-m}) p)]$.
- $(T_{B,m}^{(2)} f)(x, p) := \mathbb{E}_\mu [f(e_m, B^{(m)} p)]$.

$T_{B,m}^{(2)}$ maps $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ onto the space $\mathcal{H}_\alpha(\mathbb{P}(\mathbb{R}^m))$ of α -Hölder continuous functions, constant in x . In particular $T_{B,m}^{(2)} : \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m)) \rightarrow C(\mathbb{P}(\mathbb{R}^m))$ is a compact transformation.

■ $(T_B^{(3)} f)(x, p) := \int f d\nu_B$, where ν_B is any K_B -stationary measure. $T_B^{(3)}$ maps $L^\infty(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ onto the space of constant functions. In particular the linear transformation $T_B^{(3)} : \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m)) \rightarrow C(\mathbb{P}(\mathbb{R}^m))$ has rank 1.

We claim that for all $B \in \mathcal{V}$ and all $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ with $0 < \alpha \leq \alpha_0$, for all $n, m \in \mathbb{N}$ with $n \geq m$, and all $(x, q) \in \Sigma \times \mathbb{P}(\mathbb{R}^m)$,

- (1) $|(T_{B,n}^{(0)}f)(x, q) - (T_{B,n,m}^{(1)}f)(x, q)| \leq C \sigma^m \|f\|_\alpha.$
- (2) $|(T_{B,n,m}^{(1)}f)(x, q) - (T_{B,m}^{(2)}f)(q)| \leq C \sigma^{n-m} \|f\|_\alpha.$
- (3) $|(T_{B,m}^{(2)}f)(q) - (T_{B,n}^{(2)}f)(q)| \leq C \sigma^m \|f\|_\alpha.$

We will conclude the proposition before proving these three claims. Setting $n = 2m$ in (1) and (2), and $n = \ell$ in (3), for all $\ell \geq m$, $B \in \mathcal{V}$ and $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ with $0 < \alpha \leq \alpha_0$,

$$\|Q_B^{2m}f - T_{B,\ell}^{(2)}f\|_\infty \leq 3C \sigma^m \|f\|_\alpha. \quad (19)$$

The sequence $\{T_{B,\ell}^{(2)}f\}_{\ell \geq 0}$ is relatively compact in $C(\mathbb{P}(\mathbb{R}^m))$. Hence the set S_f of its sublimits in $(C(\mathbb{P}(\mathbb{R}^m)), \|\cdot\|_\infty)$ is non-empty. Take any $g \in S_f$ and any K_B -stationary probability measure ν_B . We claim that $g = \langle f, \nu_B \rangle \mathbf{1}$.

From (19) we have for all $m \in \mathbb{N}$,

$$\|Q_B^{2m}f - g\|_\infty \leq 3C \sigma^m \|f\|_\alpha.$$

On the other hand, since $v_\alpha(Q_B^{2m}f) \leq C \sigma^{2m} \|f\|_\alpha$, we get $v_\alpha(g) = 0$, which implies that g is constant. But $\langle Q_B^{2m}f, \nu_B \rangle = \langle f, (Q_B^{2m})^* \nu_B \rangle = \langle f, \nu_B \rangle$ implies that $\langle g, \nu_B \rangle = \langle f, \nu_B \rangle$. Therefore $g = \langle f, \nu_B \rangle \mathbf{1}$, and also

$$\|Q_B^{2m}f - \langle f, \nu_B \rangle \mathbf{1}\|_\infty \leq 3C \sigma^m \|f\|_\alpha \quad \forall m \in \mathbb{N}.$$

This concludes the proof.

To finish we still have to prove the three claims:

Claim (1): Denote by \mathcal{F}_n the sub σ -field generated by the random variables e_1, \dots, e_n . Note that for any random variable $f : X \rightarrow \mathbb{C}$

$$\mathbb{E}_x(f) = \mathbb{E}_x \{ \mathbb{E}_{e_n}(f | \mathcal{F}_n) \}.$$

Then, using this fact we have,

$$\begin{aligned} |(T_{B,n}^{(0)}f - T_{B,n,m}^{(1)}f)(x, q)| &= |\mathbb{E}_x [f(e_n, B^{(n)}q) - f(e_n, (B^{(m)} \circ T^{n-m})q)]| \\ &\leq \mathbb{E}_x [|f(e_n, (B^{(m)} \circ T^{n-m})B^{(n-m)}q) - f(e_n, (B^{(m)} \circ T^{n-m})q)|] \\ &\leq \|f\|_\alpha \mathbb{E}_x \left[\delta((B^{(m)} \circ T^{n-m})B^{(n-m)}q, (B^{(m)} \circ T^{n-m})q)^\alpha \right] \\ &\leq \|f\|_\alpha \mathbb{E}_x \left\{ \mathbb{E}_{e_{n-m}} \left[\delta((B^{(m)} \circ T^{n-m})B^{(n-m)}q, (B^{(m)} \circ T^{n-m})q)^\alpha | \mathcal{F}_{n-m} \right] \right\} \\ &\leq \|f\|_\alpha \sup_{x,p,q} \mathbb{E}_x \left[\delta(B^{(m)}p, B^{(m)}q)^\alpha \right] \\ &\leq \|f\|_\alpha \sup_{x,p \neq q} \mathbb{E}_x \left[\left(\frac{\delta(B^{(m)}p, B^{(m)}q)}{\delta(p, q)} \right)^\alpha \right] = \|f\|_\alpha \kappa_\alpha^m(B) \leq \|f\|_\alpha C \sigma^m. \end{aligned}$$

Claim (2): Defining $\varphi_{m,q}(x) := \mathbb{E}_x [f(e_m, B^{(m)}q)]$, because (K, μ) is strongly mixing on $L^\infty(\Sigma)$ we have

$$\begin{aligned} |(T_{B,n,m}^{(1)} f - T_{B,m}^{(2)} f)(x, q)| &= |\mathbb{E}_x [f(e_n, (B^{(m)} \circ T^{n-m})q)] - \mathbb{E}_\mu [f(e_m, B^{(m)}q)]| \\ &= |\mathbb{E}_x \{ \mathbb{E}_{e_{n-m}} [f(e_n, (B^{(m)} \circ T^{n-m})q)] \} - \mathbb{E}_\mu [f(e_m, B^{(m)}q)]| \\ &= |\mathbb{E}_x [\varphi_{m,q}(e_{n-m})] - \int \varphi_{m,q} d\mu| = |Q_K^{n-m} \varphi_{m,q} - \int \varphi_{m,q} d\mu \mathbf{1}| \leq C \sigma^{n-m}. \end{aligned}$$

Claim (3): Because μ is K -stationary,

$$\begin{aligned} |(T_{B,n,m}^{(2)} f)(q) - (T_{B,n}^{(2)} f)(q)| &= |\mathbb{E}_\mu [f(e_m, B^{(m)}q)] - \mathbb{E}_\mu [f(e_n, B^{(n)}q)]| \\ &= |\mathbb{E}_\mu [f(e_m, B^{(m)}q)] - \mathbb{E}_\mu [f(e_m, B^{(m)} B^{(n-m)}q)]| \\ &\leq \mathbb{E}_\mu [|f(e_m, B^{(m)}q) - f(e_m, B^{(m)} B^{(n-m)}q)|] \\ &\leq \|f\|_\alpha \mathbb{E}_\mu [\delta(B^{(m)}q, B^{(m)} B^{(n-m)}q)^\alpha] \\ &\leq \|f\|_\alpha \mathbb{E}_\mu \{ \mathbb{E}_{e_{n-m}} [\delta(B^{(m)}q, B^{(m)} B^{(n-m)}q)^\alpha] \} \\ &\leq \|f\|_\alpha \sup_{x,p,q} \mathbb{E}_x [\delta(B^{(m)}q, B^{(m)}p)^\alpha] \\ &\leq \|f\|_\alpha \sup_{x,p \neq q} \mathbb{E}_x \left[\left(\frac{\delta(B^{(m)}q, B^{(m)}p)}{\delta(q,p)} \right)^\alpha \right] = \|f\|_\alpha \kappa_\alpha^m(B) \leq \|f\|_\alpha C \sigma^m. \end{aligned}$$

□

Corollary 3.13. *Given $A \in \mathcal{B}_m^\infty(K)$ such that $(K_A, \mu_A, \xi_A) \in \mathfrak{X}$, the kernel K_A on the product space $\Sigma \times \Sigma \times \mathbb{P}(\mathbb{R}^m)$ has a unique stationary measure.*

Proof. In the proof of Proposition 3.12 we have shown that given a function $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$, if we denote by S_f the set of sublimits of $\{T_{B,\ell}^{(2)} f\}_{\ell \geq 0}$, then $S_f = \{\langle f, \nu_B \rangle \mathbf{1}\}$ for any K_B -stationary measure ν_B .

Hence, given any other K_B -stationary measure μ_B , and $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$, we have $\langle f, \nu_B \rangle = \langle f, \mu_B \rangle$. Since $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ is dense in $L^\infty(\Sigma \times \mathbb{P}(\mathbb{R}^m))$, it follows that $\nu_B = \mu_B$. □

The Laplace-Markov operator $Q_{A,z}$ of the observed Markov system (K_A, μ_A, ξ_A) is given by

$$(Q_{A,z} f)(x, y, p) = \int_\Sigma f(y, z, A(y, z)p) \|A(y, z)\|^z K(y, dz). \quad (20)$$

Like the Markov operator Q_A defined in (17), the Laplace-Markov operator $Q_{A,z}$ leaves invariant the subspaces $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$, for all small enough $\alpha > 0$. Choose $0 < \alpha_1 < \alpha_0 \leq 1$ according to proposition 3.11.

Assumption (A3) is automatically satisfied because $\|A\|_\infty < \infty$ and $\|A^{-1}\|_\infty < \infty$ which imply that $\xi_A \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ for all $\alpha > 0$. Note that $Q_{A,z} = Q_A \circ D_{e^{z\xi_A}}$, where $D_{e^{z\xi_A}}$ denotes the multiplication operator by $e^{z\xi_A}$. This is a bounded operator because $\mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$ is a Banach algebra containing the function $e^{z\xi_A}$.

Finally the next lemma proves (A4).

Lemma 3.14. *Given $A, B \in \mathcal{B}_m^\infty(K)$ and $b > 0$, there is a constant $C_2 > 0$ such that for all $f \in \mathcal{H}_\alpha(\Sigma \times \mathbb{P}(\mathbb{R}^m))$, and all $z \in \mathbb{C}$ such that $\operatorname{Re} z \leq b$,*

$$\|Q_{A,z}f - Q_{B,z}f\|_\infty \leq C_2 d_\infty(A, B)^\alpha \|f\|_\alpha.$$

Proof. A simple computation shows that for all $z \in \mathbb{C}$ with $\operatorname{Re} z \leq b$, and all $A, B \in \operatorname{GL}(d, \mathbb{R})$,

$$\| \|Ap\|^z - \|Bp\|^z \| \leq b \max\{\|A\|^{b-1}, \|B\|^{b-1}\} \|A - B\|.$$

Hence

$$\begin{aligned} |(Q_{A,z}f - Q_{B,z}f)(x, p)| &\leq \mathbb{E}_x [\| \|Ap\|^z f(e_1, Ap) - \|Bp\|^z f(e_1, Bp) \|] \\ &\leq \|f\|_\infty \mathbb{E}_x [\| \|Ap\|^z - \|Bp\|^z \|] + \|B\|_\infty^b \mathbb{E}_x [|f(e_1, Ap) - f(e_1, Bp)|] \\ &\leq b \max\{\|A\|_\infty^{b-1}, \|B\|_\infty^{b-1}\} \|A - B\|_\infty \|f\|_\infty + \|B\|_\infty^b v_\alpha(f) \mathbb{E}_x [\delta(Ap, Bp)^\alpha] \\ &\leq b \max\{\|A\|_\infty^{b-1}, \|B\|_\infty^{b-1}\} \|A - B\|_\infty \|f\|_\infty \\ &\quad + \|B\|_\infty^b v_\alpha(f) \|A - B\|_\infty^\alpha \leq C_2 \|f\|_\alpha d_\infty(A, B)^\alpha, \end{aligned}$$

where $C_2 = \max\{\|B\|_\infty^b, b\|B\|_\infty^{b-1}, b\|A\|_\infty^{b-1}\}$. \square

Proof of Theorem 1.3. The space of observed Markov systems \mathfrak{X} satisfies all assumptions (A1)-(A4). Hence, by Theorem 2.1, there exists a neighborhood \mathcal{V} of $(K_A, \mu_A, \xi_A) \in \mathfrak{X}$, which we identify with a neighborhood of $A \in \mathcal{B}_m^\infty(K)$, and there are constants $\varepsilon_0, C, h > 0$ such that for all $B \in \mathcal{V}$, $0 < \varepsilon < \varepsilon_0$, $(x, p) \in \Sigma \times \mathbb{P}(\mathbb{R}^m)$ and $n \in \mathbb{N}$,

$$\mathbb{P}_x \left[\left| \frac{1}{n} \log \|B^{(n)} p\| - L_1(B, \mu) \right| \geq \varepsilon \right] \leq C e^{-\frac{\varepsilon^2}{2h} n}.$$

Integrating w.r.t. μ we get for all $p \in \mathbb{P}(\mathbb{R}^m)$,

$$\mathbb{P}_\mu \left[\left| \frac{1}{n} \log \|B^{(n)} p\| - L_1(B, \mu) \right| \geq \varepsilon \right] \leq C e^{-\frac{\varepsilon^2}{2h} n}.$$

Choose the canonical basis $\{e_1, \dots, e_d\}$ of \mathbb{R}^m and consider the following norm $\|\cdot\|'$ on the space of matrices $\operatorname{Mat}_d(\mathbb{R})$, $\|M\|' := \max_{1 \leq j \leq d} \|M e_j\|$. Since this norm is equivalent to the operator norm, for all $B \in \mathcal{V}$, $p \in \mathbb{P}(\mathbb{R}^m)$ and $n \in \mathbb{N}$,

$$\|B^{(n)} p\| \leq \|B^{(n)}\| \lesssim \|B^{(n)}\|' = \max_{1 \leq j \leq d} \|B^{(n)} e_j\|.$$

Thus a simple comparison of the deviation sets gives

$$\mathbb{P}_\mu \left[\left| \frac{1}{n} \log \|B^{(n)}\| - L_1(B, \mu) \right| \geq \varepsilon \right] \lesssim e^{-\frac{\varepsilon^2}{2h} n}$$

for all $B \in \mathcal{V}$, $0 < \varepsilon < \varepsilon_0$ and $n \in \mathbb{N}$. \square

4. DERIVING CONTINUITY OF THE LYAPUNOV EXPONENTS

In this last section we use the LDT estimates (theorems 1.2 and 1.3) to derive the continuity of the Lyapunov exponents and of the Osledeets's filtration / decomposition. We give some simple generalizations of the continuity results and explain the method's limitations regarding the continuity of the LE in the reducible case.

4.1. Proof of the continuity.

Proof of Theorem 1.1. Let (K, μ) be a strongly mixing Markov system, and consider the associated Markov shift (X, \mathbb{P}_μ, T) .

The collection $\mathcal{C} = \{(\mathcal{I}_m^\infty(K), d_\infty)\}_{m \in \mathbb{N}}$ is a space of measurable cocycles in the sense of [7, Definition 1.8] (see also [4, Definition 1.1]). We are going to apply the abstract continuity theorem (ACT) [7, Theorem 1.6] (see also ([4, Theorem 1.1] and [6, Theorems 3.2 and 3.3]) to this space of totally irreducible cocycles over (X, \mathbb{P}_μ, T) .

Consider the space of LDT parameters $\mathcal{P} = \mathbb{N} \times \mathcal{E} \times \mathcal{J}$, where \mathcal{E} is the set of constant deviation functions $\underline{\epsilon}(t) \equiv \epsilon$, $0 < \epsilon < 1$, and we use the set of exponential functions $\mathcal{J} = \{\underline{c}(t) \equiv M e^{-ct} : M < \infty, c > 0\}$ to measure the deviation sets.

Define Ξ to be the set of observables $\xi : X \rightarrow \mathbb{R}$ which depend only on finitely many coordinates. Finally, take $p = \infty$.

We now check the four assumptions of the ACT.

1. The set Ξ is compatible with all cocycles $A \in \mathcal{B}_m^\infty(K)$, because for any set $F \in \mathcal{F}_N(A)$ its indicator function $\mathbf{1}_F$ depends only on finitely many coordinates, i.e., $\mathbf{1}_F \in \Xi$.

2. Given an observable $\xi \in \Xi$ there exists $p \in \mathbb{N}$ such that $\xi \circ T^p$ depends only on negative coordinates, i.e., coordinates x_j with $-p \leq j \leq 0$. This implies that $\xi \circ T^p \in \mathcal{H}_\alpha(X^-)$. By Theorem 1.2, the observable $\xi \circ T^p$ satisfies a base-LDT estimate w.r.t. \mathcal{P} . Since $|S_n(\xi) - S_n(\xi \circ T^p)|$ converges uniformly to zero as $n \rightarrow \infty$, it follows that ξ satisfies base-LDT estimates too.

3. The L^p -boundedness assumption is automatic because $p = \infty$ and the functions A and A^{-1} are bounded.

4. Given $A \in \mathcal{I}_m^\infty(K)$ such that $L_1(A) > L_2(A)$, by Theorem 1.3 the cocycle A satisfies uniform fiber-LDT estimates w.r.t. \mathcal{P} .

A simple computation shows that the modulus of continuity associated to the choice of deviation function sets \mathcal{E} and \mathcal{J} above corresponds to Hölder continuity. Hence, this theorem follows from the conclusions of the ACT. \square

4.2. Some generalizations. Consider a compact metric space Σ .

A *Markov kernel* of order $p \in \mathbb{N}$ on Σ is a map $K : \Sigma^p \rightarrow \text{Prob}(\Sigma)$ that assigns a probability measure $K(x_0, \dots, x_{p-1}, dy)$ on Σ to each tuple $(x_0, \dots, x_{p-1}) \in \Sigma^p$. The concept of Markov kernel in Definition 1.1 corresponds to a Markov kernel of order $p = 1$.

Any Markov kernel K of order p on Σ determines the following Markov kernel \hat{K} of order 1 on the product space Σ^p ,

$$\hat{K}(x_0, \dots, x_{p-1}) := \int_{\Sigma} \delta_{(x_1, \dots, x_p)} K(x_0, \dots, x_{p-1}, dx_p) .$$

A probability measure μ on Σ^p is said to be *K-stationary* when it is \hat{K} -stationary. We call a Markov system of order p any pair (K, μ) , where K is a Markov kernel of order p on Σ , and μ is a K -stationary probability on Σ^p . We say that (K, μ) is strongly irreducible when (\hat{K}, μ) is a strongly irreducible Markov system on Σ^p .

Given a Markov system (K, μ) of order p , let $\hat{\mathbb{P}}_\mu$ denote the Kolmogorov extension of (K, μ) on the space of sequences $\hat{X} := (\Sigma^p)^\mathbb{Z}$. Then, letting $\hat{T} : \hat{X} \rightarrow \hat{X}$ denote the shift homeomorphism, the triple $(\hat{X}, \hat{\mathbb{P}}_\mu, \hat{T})$ is a Markov shift.

Let $X := \Sigma^\mathbb{Z}$ and consider the maps

- $\psi : X \rightarrow \hat{X}, \psi\{x_n\}_{n \in \mathbb{Z}} = \{(x_n, \dots, x_{n+p-1})\}_{n \in \mathbb{Z}},$
- $\pi : \hat{X} \rightarrow X, \pi\{(x_{0,n}, \dots, x_{p-1,n})\}_{n \in \mathbb{Z}} = \{x_{0,n}\}_{n \in \mathbb{Z}},$

which satisfy $\pi \circ \psi = \text{id}_X$.

Defining $\mathbb{P}_\mu := \pi_* \hat{\mathbb{P}}_\mu$, these maps are bimeasurable isomorphisms conjugating the shifts on $(\hat{X}, \hat{\mathbb{P}}_\mu)$ and (X, \mathbb{P}_μ) , where the measure \mathbb{P}_μ is invariant under the shift $T : X \rightarrow X$. The triple (X, \mathbb{P}_μ, T) is called a Markov shift of order p .

Consider now the space $\mathcal{B}_m^\infty(K, p)$ of measurable functions $A : X \rightarrow \text{GL}(m, \mathbb{R})$ which depend only on the coordinates $(x_0, \dots, x_p) \in \Sigma^{p+1}$ with $\|A\|_\infty < \infty$ and $\|A^{-1}\|_\infty < \infty$. Note that the iterates of A are

$$A^{(n)}(x) = A(x_{n-1}, \dots, x_{n-1+p}) \dots A(x_1, \dots, x_{1+p}) A(x_0, \dots, x_p) .$$

We identify $\mathcal{B}_m^\infty(K, p)$ as a space of functions $A : \Sigma^{p+1} \rightarrow \text{GL}(m, \mathbb{R})$. Each such function determines a locally constant cocycle over the Markov shift (X, \mathbb{P}_μ, T) .

Given $A \in \mathcal{B}_m^\infty(K, p)$, we define $\hat{A} : \Sigma^p \times \Sigma^p \rightarrow \text{GL}(m, \mathbb{R})$

$$\hat{A}((x_0, \dots, x_{p-1}), (y_0, \dots, y_{p-1})) := A(x_0, \dots, x_{p-1}, y_{p-1}) .$$

Identifying \hat{A} with a function $\hat{A} : \hat{X} \rightarrow \text{GL}(m, \mathbb{R})$ we have $\hat{A} \circ \psi = A$. Hence the cocycles (\hat{T}, \hat{A}) and (T, A) are conjugated.

The cocycle (T, A) over the Markov shift (X, \mathbb{P}_μ, T) will be called a *random Markov cocycle* of order p .

Define $\mathcal{I}_m^\infty(K, p)$ to be the subspace of *totally irreducible* cocycles $A \in \mathcal{B}_m^\infty(K, p)$, i.e., the subspace of cocycles A such that \hat{A} is totally irreducible over $(\hat{X}, \hat{\mathbb{P}}_\mu, \hat{T})$.

From these considerations and Theorem 1.1 we obtain the following result.

Theorem 4.1. *Let (K, μ) be a strongly mixing Markov system of order $p \in \mathbb{N}$.*

Then all Lyapunov exponents $L_j : \mathcal{I}_m^\infty(K, p) \rightarrow \mathbb{R}$, with $1 \leq j \leq m$, the Oseledets filtration $F : \mathcal{I}_m^\infty(K, p) \rightarrow \mathfrak{F}(X, \mathbb{R}^m)$, and the Oseledets decomposition $E : \mathcal{I}_m^\infty(K, p) \rightarrow \mathfrak{D}(X, \mathbb{R}^m)$, are continuous functions of the cocycle $A \in \mathcal{I}_m^\infty(K, p)$.

Moreover, if $A \in \mathcal{I}_m^\infty(K, p)$ has a τ -gap pattern then the functions Λ^τ , F^τ and E^τ are Hölder continuous in a neighborhood of A .

In particular, all conclusions above on the continuity of the LE, the Oseledets filtration, and the Oseledets decomposition, apply to irreducible and locally constant cocycles over strongly mixing Markov and Bernoulli shifts.

The abstract setting developed in section 2 is general enough to deal with cocycle having singularities, i.e., points $x \in X$ where the matrix $A(x)$ is singular. Consider the family of spaces $\mathcal{B}_m^a(K)$, with $0 < a < \infty$, consisting of all bounded measurable functions $A : \Sigma \times \Sigma \rightarrow \text{GL}(m, \mathbb{R})$ such that for some $C > 0$ and all $x \in \Sigma$,

$$\eta_A^a(x) := \int_{\Sigma} \|A(x, y)^{-1}\|^a K(x, dy) \leq C .$$

Equip this space with the distance

$$d_a(A, B) := \|A - B\|_\infty + \|\eta_A^a - \eta_B^a\|_\infty .$$

The collection $\mathcal{C} = \{(\mathcal{B}_m^a(K), d_a)\}_{m \in \mathbb{N}}$ is not a space of measurable cocycles, because (2) of [7, Definition 1.8] (see also [4, Definition 1.1])

fails. However, both the uniform fiber-LDT estimates and the continuity statements about the LE can be extended to the spaces $\mathcal{I}_m^a(K)$ of totally irreducible cocycles in $\mathcal{B}_m^a(K)$. More precisely, it can be proved that Theorem 1.3 holds for all $a \geq 4$, and Theorem 1.1 holds for all $a \geq 4m$.

4.3. Method limitations. We need the irreducibility assumption in order to prove uniform fiber LDT estimates in Theorem 1.3. The proof exploits the fact that for irreducible cocycles there is some Banach algebra of measurable functions, *independent* of the cocycle, where the associated Laplace-Markov operators act as quasi-compact and simple operators (see Proposition 3.12). For reducible cocycles this fact may still be true, and it could eventually lead to fiber LDT estimates. However, the Banach algebra would have to be tailored to the cocycle, and hence the scheme of proof presented here would not provide the required uniformity.

Acknowledgments. Both authors would like to acknowledge fruitful conversations they had with A. Baraviera, G. Del Magno and J. P. Gaivão about large deviation estimates for random cocycles.

The first author was supported by Fundação para a Ciência e a Tecnologia, UID/MAT/04561/2013.

The second author was supported by the Norwegian Research Council project no. 213638, "Discrete Models in Mathematical Analysis". He is also grateful to the University of Lisbon and to the Institut Mittag-Leffler (through its "Research in Peace" program) for their hospitality during the summer of 2013, when this project began.

REFERENCES

1. Philippe Bougerol, *Théorèmes limite pour les systèmes linéaires à coefficients markoviens*, Probab. Theory Related Fields **78** (1988), no. 2, 193–221. MR 945109 (89i:60122)
2. Philippe Bougerol and Jean Lacroix, *Products of random matrices with applications to Schrödinger operators*, Progress in Probability and Statistics, vol. 8, Birkhäuser Boston, Inc., Boston, MA, 1985. MR 886674 (88f:60013)
3. J. L. Doob, *Stochastic processes*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1990, Reprint of the 1953 original, A Wiley-Interscience Publication. MR 1038526 (91d:60002)
4. Pedro Duarte and Silviu Klein, *An abstract continuity theorem of the Lyapunov exponents*, preprint (2015), 1–32.
5. ———, *The Avalanche Principle and other estimates on Grassmann manifolds*, preprint (2015), 1–49.
6. ———, *Continuity of the Oseledets decomposition*, preprint (2015), 1–32.

7. ———, *Lyapunov exponents of linear cocycles, continuity via large deviations*, Atlantis Studies in Dynamical Systems, Atlantis Press, to appear in 2016.
8. H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Statist. **31** (1960), 457–469. MR 0121828 (22 #12558)
9. Harry Furstenberg, *Noncommuting random products*, Trans. Amer. Math. Soc. **108** (1963), 377–428. MR 0163345 (29 #648)
10. Y. Guivarc’h and A. Raugi, *Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence*, Z. Wahrsch. Verw. Gebiete **69** (1985), no. 2, 187–242. MR 779457 (86h:60126)
11. Hubert Hennion and Loïc Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Lecture Notes in Mathematics, vol. 1766, Springer-Verlag, Berlin, 2001. MR 1862393 (2002h:60146)
12. Alexander S. Kechries, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.
13. S G Krein and Yu I Petunin, *Scales of banach spaces*, Russian Mathematical Surveys **21** (1966), no. 2, 85.
14. Émile Le Page, *Théorèmes limites pour les produits de matrices aléatoires*, Probability measures on groups (Oberwolfach, 1981), Lecture Notes in Math., vol. 928, Springer, Berlin-New York, 1982, pp. 258–303. MR 669072 (84d:60012)
15. Eugene Lukacs, *Characteristic functions*, Hafner Publishing Co., New York, 1970, Second edition, revised and enlarged. MR 0346874 (49 #11595)
16. S. V. Nagaev, *Some limit theorems for stationary Markov chains*, Teor. Veroyatnost. i Primenen. **2** (1957), 389–416. MR 0094846 (20 #1355)
17. Karl Petersen, *Ergodic theory*, Cambridge Studies in Advanced Mathematics, vol. 2, Cambridge University Press, Cambridge, 1983. MR 833286 (87i:28002)
18. A. Raugi and J. Rosenberg, *Fonctions harmoniques et théorèmes limites pour les marches aléatoires sur les groupes*, Bulletin de la Société mathématique de France. Mémoire, no. no. 54, Société mathématique de France, 1977.
19. F. Riesz and B. Szőkefalvi-Nagy, *Functional analysis*, Ungar, 1955.
20. V. N. Tutubalin, *Limit theorems for a product of random matrices*, Teor. Veroyatnost. i Primenen. **10** (1965), 19–32. MR 0175169 (30 #5354)

DEPARTAMENTO DE MATEMÁTICA AND CMAFIO, FACULDADE DE CIÊNCIAS,
UNIVERSIDADE DE LISBOA, PORTUGAL
E-mail address: pmduarte@fc.ul.pt

DEPARTMENT OF MATHEMATICAL SCIENCES, NORWEGIAN UNIVERSITY OF
SCIENCE AND TECHNOLOGY (NTNU), TRONDHEIM, NORWAY, AND IMAR, BUCHAREST,
ROMANIA
E-mail address: silviu.klein@math.ntnu.no